DOCUMENTO DE TRABAJO 9507

DETECTION OF CHAOS IN TIME SERIES. APPLICATION TO SPANISH SEA SWELL.

Eugenia Mera
Manuel Moran
José Manuel Rey

FACULTAD DE CIENCIAS ECONOMICAS Y EMPRESARIALES
UNIVERSIDAD COMPLUTENSE
Campus de Somosaguas 28223 MADRID

DETECTION OF CHAOS IN TIME SERIES. APPLICATION TO SPANISH SEA SWELL.

Eugenia Mera Manuel Morán José Manuel Rey

Universidad Complutense
Madrid
March, 1995

Memoria del proyecto de investigación

"DETECCION DE CAOS DETERMINISTA EN EL OLEAJE MARINO"

firmado entre la Universidad Complutense y la entidad pública *Puertos del Estado*

Director del proyecto:
Manuel Morán

Autores:
Eugenia Mera Manuel Morán José Manuel Rey

Marzo, 1995

Abstract

En esta memoria se revisan tres herramientas complementarias para el *Análisis No-Lineal de Series Temporales*: dimension fractal, entropía, y exponentes de Liapunov. Se hace una revisión crítica del estado de la cuestión en el área, señalando el fundamento teórico de las técnicas aplicadas cuando éste existe, el campo de aplicación admitido por los especialistas cuando no existe base teórica, las limitaciones, etc. En ocasiones se aportan pruebas rigurosas, o razonamientos heurísticos, nuevos en la literatura.

Se seleccionan, perfeccionan, y traducen en códigos FORTRAN algoritmos que permiten cuantificar las magnitudes arriba mencionadas tanto para una serie monitorizada en un sistema dinámico determinista como en una realización de un proceso estocástico estacionario. Ello posibilita el tratamiento de datos provenientes de un amplio espectro de áreas, tanto de las ciencias de la naturaleza como sociales o económicas.

El análisis numérico propuesto permite poner de manifiesto características relevantes del proceso estudiado; en particular permite detectar la existencia de una dinámica caótica como mecanismo generador de los datos. Se introducen dos nuevos tests, uno basado en la teoría de exponentes de Liapunov, capaz de detectar la existencia de dinámica caótica diferenciable en series experimentales; otro capaz de cuantificar la sensibilidad a condiciones iniciales de una dinámica no diferenciable.

Se aplican estas técnicas, combinadas con técnicas de análisis espectral, a series temporales de altura significativa de ola registradas por boyas escalares emplazadas en la costa del mar Cantábrico. Como consecuencia de este análisis, y a la vista de los datos de los que se dispone, se concluye que el número de variables capaces de explicar la dinámica de las series de altura de ola es elevado, con una cota inferior del orden de quince, y presumiblemente bastante mayor. También se infieren estimaciones del grado de aleatoriedad de los datos (medidos según la entropía), que resulta ser elevado en comparación a dinámicas caóticas como la del flujo de Lorenz o la aplicación de Henon. Se concluye la no adecuación de los métodos de predicción y filtrado lineales para el tratamiento de estos datos. También se constatan notables propiedades de ergodicidad de las series bajo estudio, que permiten abrigar esperanzas de éxito en su futuro estudio de estos procesos mediante nuevas técnicas no-lineales en activo desarrollo.

Contents

List of Figures

List of Tables

Preamble

The research presented in this report has been devised with the aim of determining whether there exists chaotic behaviour in significant wave height time series –SWH–series from now onwards–, read by buoys moored on the Spanish northern coast. The main techniques reviewed in the report have been implemented by our research team as FORTRAN codes which have been tested in classical dynamical systems as well as in stochastic processes, and then applied to the analysis of the mentioned data. The results of such analysis are reported throughout the text, after the exposition of their significance and theoretical basis.

Many tools in use in this new area of research are founded on theoretical results that can only be found in specialized research papers; others are used by experimentalist researchers who only care that the new methods are based on sound ideas supported with some empirical evidence, but without definite theoretical proofs. We have developed a systematic work of reviewing the extensive literature. We report on which methods have a solid theoretical basis, giving in many cases new proofs or heuristic arguments for them, and in any case the references where the available proofs can be found. We report also on the methods which are most extensively accepted while not being properly proved. We discuss its significance and scope, and in some cases we give proofs for them.

The source of some of the new ideas proposed by us is an empirical one: they are connected with the attempt at explanation of the results obtained for the SWH–series. We actually think that we are able to give a definite answer to the question motivating this research. This is an answer in the negative: with the recorded data in view, the hypothesis of existence of a low–dimensional chaotic dynamics may be rejected. The methods applied to obtain such an answer yield besides valuable information about certain features of the data. They also have methodological consequences on the analysis of time series: we propose a new test for detection of chaos, based on the computation of Liapunov spectrum and various modifications

to the algorithms in use. The hurried reader can find a summarized information of these and other results in chapter five, devoted to the conclusions. An appendix **C**, collecting the figures and tables referred to throughout this report, has been written in such a way that it can be read independently of the rest of the text. This introduction, together with appendix **C** and chapter five, gives a panoramic view of the work that we have developed here.

The methods and concepts reviewed in this report are devised to *detect the existence of chaotic dynamics in a time series.* They can be considered as a part of an emerging branch of applied mathematics called *non–linear analysis of time series.* In chapter one, a motivation for the detection of chaos is presented, the *Strange Attractor Hypothesis* (SAH) is formally formulated, and two basic tools for the non–linear analysis of time series, *embedology* and *ergodic theory*, are briefly reviewed.

Chapter two describes the techniques for the measurement of the dimensional concepts which can be applied to the analysis of time series, and focuses on the most useful of them, namely the *correlation dimension*, which can be computed through the Grassberger–Procaccia algorithm. We describe the obtention and significance of the SWH–series, propose a specific methodology for its numerical treatment, and apply to the data the mentioned algorithm, together with certain tests devised for guaranteeing the robustness of the analysis.

The K_2–entropy provides a lower bound for the randomness of a dynamics measured in terms of entropy. This permits in some cases to rule out low–dimensional chaos. The theoretical concepts of *Kolmogorov–Sinai entropy* and *K_2–entropy*, along with the practical algorithm for its computation, are reviewed in chapter three, and the analysis of these parameters combined with methods of spectral theory are carried out over the SWH–series.

Chapter four deals with the computation of *Liapunov exponents*, a tool recently developed for the statistical analysis of smooth dynamics that can also be applied to some stochastic processes. We describe the theoretic background of this tool and its main applications. We describe in detail an algorithm for its computation, based upon some modifications of the algorithm by J.P. Eckmann and D.

Ruelle [Eckmann & Ruelle,85]. These modifications, though very simple, improve the computation of the Liapunov exponents, specially in noisy time series. The results obtained for the SWH–series, as well as for some stochastic processes, suggest a natural and, seemingly, rather conclusive test for the detection of low–dimensional chaotic dynamics.

We summarize the main results obtained in this research in the brief chapter five, devoted to the conclusions. Lastly, we review in several appendices complementary aspects which may help to understand the scope of the concepts in the previous chapters, and also alternative tests and theoretic tools which are in use in non–linear analysis of time series.

The research here reported has been sponsored by PUERTOS DEL ESTADO (formerly *Dirección General de Puertos del Ministerio de Obras Públicas y Transportes*). We are grateful to José Conde, who devised the topic of this research project and trusted it to our team, and to Antonio Ruiz de Elvira for his support to the further development of the research as director of the section of *Climatología Marítima del Ente Puertos del Estado.* The collaboration with both José Carlos Nieto and Marta Alfonso, members of the technical staff of the mentioned entity, has gone far beyond a technical support: it has created a dynamics of true team work and of stimulating scientific discussions which has produced results unthinkable for a more conventional cooperation, as proved by several joint research papers that are in course of production at the moment of writing this report. In particular, we would like to thank José Carlos and Marta for their excellent supporting work on the spectral techniques referred to in this report, namely filtering the spectrum and randomizing the Fourier phases of the SWH series.

We are grateful to all our colleagues in the department of Análisis Económico at the Universidad Complutense, especially to its chairman Carmen Carrera, who has facilitated our work with a constant and effective support, and to professors Antonio Rodrigo and Mercedes Vázquez who have contributed with many discussions to this research.

Professors Miguel Angel Martín and Miguel Reyes have also spent many hours in common discussions held in the sessions of the seminar of non–linear analysis

and ergodic theory of the department, in the Conferences given by them and by the members of our team, and on many other occasions. We are also grateful to the staff of the Centro de Cálculo de Somosaguas, which have suffered the always growing needs on CPU and memory space demanded by this research to a Convex which was born overwhelmed from the very first day by the task it had to accomplish. Lastly we are grateful to Miguel de Guzmán, who initiated the director of this research team into Fractal Geometry, a field that nowadays attracts an increasing interest in our country as a result of that pioneering work started about ten years ago by Miguel de Guzmán in the seminar of Geometric Measure Theory of the Faculty of Mathematics of the Universidad Complutense.

Eugenia Mera, Manuel Morán, José Manuel Rey.

Madrid, March 1995.

Chapter 1

An Outline of Detection of Chaos in Time Series

1.1 Why Chaotic Dynamics should be Tested in Time Series?. The Strange Attractor Hypothesis.

A deterministic (discrete) dynamical system is a system steered by an equation of the form

$$x_{k+1} = f(x_k), \tag{1.1}$$

where x_{k+1} is the vector describing the state of the system after $k+1$ units of time, and f is the evolution law, *determining* precisely the state of the system at any period of time from the knowledge of its state in the previous period. Thus, in a deterministic dynamics, the values $f(x_0), f^2(x_0) = f(f(x_0)), \ldots, f^k(x_0)$, recursively permit the exact prediction of the state of the system at any future period of time from an exact knowledge of its initial state x_0 (called initial condition). Recently much research work has focused on the topic of detecting the existence of such an evolution law in empirically recorded time series. This interest started from the discovery that simple deterministic dynamical systems can generate time series with a degree of randomness as high as those generated by purely random stochastic

processes. Let us examine the following

1.1.1 Example (Saw–Tooth Dynamics).

Let $F : [0,1] \mapsto [0,1]$ be the mapping given by

$$F(x) = 2x - \text{int}(2x), \tag{1.2}$$

where $\text{int}(x)$ stands for the integer part of the real number x. The deterministic dynamical system obtained by taking this function F as the evolution law in (1.1) can generate time series as random as that generated by a fair (or weighed) coin. In order to see it suppose further that our measurement device can give the value of an observation x with one bit of accuracy, that is, it gives the value $h(x) = 0$ for $x \in [0, 1/2)$ and the value $h(x) = 1$ for $x \in [1/2, 1]$. It is easy to see that the observed sequence of values $h(F^k(x))$ can be *any sequence of zeroes and ones.* Let $\omega \in \Omega = \{0,1\}^{\mathbb{N}}$ be one of such sequences, and consider the mapping $\pi : \Omega \mapsto [0,1]$ defined by

$$\pi(\omega) = \sum_{i \in \mathbb{N}} \omega_i 2^{-i}.$$

It is easily seen that $h(F^k(\pi(\omega))) = \omega_{k+1}$. Therefore we can obtain the arbitrary sequence $\omega_1, \omega_2, \omega_3 \ldots$ as outcomes of our observation of the system $h(F^k(x))$, $k = 0, 1, 2, \ldots$ by setting $x = \pi(\omega)$. Assume now that we construct the sequence ω by tossing a fair coin. This is equivalent to regarding the set Ω of all possible sequences as the probability space of a stochastic process $\{X_i\}_{i \in \mathbb{N}}$, where X_i are independent identically distributed (i.i.d.) random variables with probability distribution given by $P(X_i = 0) = 1/2$ and $P(X_i = 1) = 1/2$. The mapping π maps Ω onto $[0,1]$, and it can be shown [Billingsley,79] that if ν is the probability measure defined in Ω by the stochastic process, the measure $\nu \circ \pi^{-1}$ induced by ν in $[0,1]$ is the standard Lebesgue measure L^1. In fact, the probability spaces (Ω, ν) and $([0,1], L^1)$ are isomorphic. This means that, from a probabilistic point of view, the outcome of an infinite sequence of tossings of a fair coin is perfectly equivalent to the observation $h(F^k(x_0))$, $k = 0, 1, 2, \ldots$, of a whole orbit of the dynamical system if the initial condition x_0 is selected according to the Lebesgue measure in the unit interval. In other words, the dynamic system can mimic exactly the tossings of a fair coin.

Weighed coins are mimicked by the same dynamical system by selecting the initial condition according to certain self–similar measures [Morán & Rey,94].

The feature of being able to generate random time series is shared by a wide family called *chaotic dynamical systems* (see [Devaney,87]). They often generate orbits that, in the stationary regime, wander over complex geometric structures called *chaotic attractors* (although a chaotic dynamics can also occur on the unit interval, as in the case of the saw–tooth mapping (1.2)). Such systems can generate smooth dynamics. They arise often as the flow of a low–dimensional dynamical system of differential equations (three is the minimum allowed number of equations) or as the discrete flows of diffeomorphism (here it is possible to find chaos in unidimensional dynamics). It should not cause much surprise that they are able to fool the standard tests for randomness since they generate genuine randomness. The current methods of analysis of random time series, such as the spectral analysis or the computation of the autocorrelation function (ACF), do not permit the detection of a deterministic generating mechanism of the analyzed series. As an example, in the case of the saw–tooth dynamics it can be proved [Sakai & Tokumaru,80] that the ACF–coefficients are, with probability one, those of a first–order autoregresive process. The concern of the research area mentioned above is the developing of techniques capable of detecting the existence of a chaotic attractor as the generating mechanism of the randomness of an observed time series.

The existence of chaotic dynamics in a given time series has happened to result in an exciting possibility for many scientists, since it opens up the perspective of the introduction of the powerful tools of analysis and prediction available in classical dynamical systems theory. But this is not the only consequence that can be drawn from examples such as the saw–tooth dynamics. These examples may be read in the opposite sense, and this reading also entails profound implications. The following question concentrates on these: does there exist in empirical sciences purely deterministic dynamics?. As we have remarked above, an exact knowledge of an initial condition x_0 gives the whole forward orbit in a deterministic dynamics. But exactly determining x_0 involves the infinite process of determination of each one of its binary digits. The isomorphism between the probability spaces (Ω, ν) and $([0,1], L^1)$ shows that there is as much randomness in determining exactly x_0 as in an infinite sequence of tossings of a fair coin. Any measurement device for observing a deter-

ministic dynamics can determine an initial condition with a given number k of bits of accuracy. In the case of the saw–tooth dynamics, this means that the outcomes of the observed system, even if we know exactly its evolution law, can be predicted at the i-th period of time within an accuracy of 2^{k-i} bites, and in consequence in k periods of time we have lost all capacity of prediction. Thus, the saw–tooth dynamics when observed by our measurement device at intervals of k–periods of time behaves like a fair die of 2^k sides. There is not any way of getting around this problem. The dual aspect of the problem thus arises: the most fruitful approach to the analysis and prediction of a deterministic dynamics of chaotic behaviour must take into account its intrinsic randomness, and thus incorporate the probabilistic point of view afforded by the classical theory of stochastic systems.

The goal is thus a double one. There emerge methods such as *embedology*, that geometrizes any time series including those generated by stochastic processes, as well as concepts (such as the *invariant measures* (see section **1.3**)) that permit deterministic dynamical systems to be regarded as stochastic processes, and stochastic processes as dynamical systems under the dynamics given by their time evolution; the *entropy* which measures the randomness of stochastic processes in terms of their predictability, and also the randomness of dynamical systems as the rate of destruction of the information about the initial conditions. All these concepts are part of the ergodic theory, a branch of mathematics focused on the analysis of dynamics in probability spaces. This covers both deterministic dynamical systems and stochastic processes. Some important tools which incorporate metric notions, such as the *sensitivity to initial conditions* or the *Liapunov exponents* (see chapter **4**) have not been sucessfully developed in stochastic processes yet, though active research is addressing these issues.

A program addressing the distinction between random noise and deterministic chaos (contaminated by some kind of noise) has been, and still is, pursued by different specialists. These include dynamical systems specialists ([Eckmann & Ruelle,85], [Ruelle,87] and references therein), researchers from a large number of applied scientific disciplines (the list is certainly uncountable, see [Grassberger et al,91], [Brock et al,92] for a picture on some areas), and in particular statisticians [Tong,90], [Bartlett,90], [Isham,93], [Jensen,93] (see also the special number from the Royal Statistical Society on its meeting on chaos [JRSS,92]).

A program on the diagnosis of chaos in time series is currently recommended at least in two cases, namely when the signal generating apparatus is suspect of being governed by some kind of deterministic equations (as in the famous example of the onset of turbulence in fluid dynamics, [Ruelle & Takens,71], [Turbulence Seminar,75]) and also when some trace of determinism seems to remain in the signal when a statistical approach is tried to explain the genesis of the series (as has been shown to occur in economic series [Brock et al,92], [Scheinkman & Lebaron,89]). However, as a consequence of the interplay between determinism and randomness sketched above, many of the methods that were initially devised for the detection of chaos are still meaningful for the analysis of any time series to which they can be applied, as we shall show throughout this report.

1.1.2 The Strange Attractor Hypothesis (SAH).

Assume that we are given a real–valued, discrete–time data recording

$$\{u_k\}_k \subset \mathbb{R}, \quad k = 1, 2, ..., N_p, \tag{1.3}$$

that we want to analyze. We now proceed to describe the basis for a deterministic explanation of time–series. Let $\Omega \subset \mathbb{R}^M$ be the phase space of a one-parameter semigroup of smooth (typically non-linear) transformations

$$\{f^t\}_{t \in T}, \quad f^t : \Omega \mapsto \Omega, \tag{1.4}$$

where $T = \mathbb{R}^+$ (continuous (semi)flow) or $T = \mathbb{N}$ (discrete-time (semi)flow). We will assume $T = \mathbb{N}$, and then understand either $f^k = f \circ f \circ \overset{(k)}{...} \circ f$ in the discrete case, or $f^k = f^{k\Delta t}$, where Δt is some positive delay, in the continuous case. Given an initial state of the system $x_0 \in \Omega$, $f^t(x_0)$ represents the state of the system after t units of time.

Assume that (1.4) is a *dissipative* flow. This means that phase–space volumes are contracted by the time evolution [Ruelle,87]. Thus it can be supposed that $f^t(\Omega) \subset \Omega$ for $t \geq 0$. The interesting case occurs when this volume compression does not happen to occur in all possible directions, but instead some directions are dilated by the system as time evolves. It is this combination of a phase-space

expansion along some directions with the global contraction that makes it possible for an unstable dynamical evolution to take place on a bounded dissipating region. When the transient behaviour vanishes, the system adopts its asymptotic regime which takes place on the following set

$$\Lambda = \bigcap_{t \geq 0} f^t(\Omega). \tag{1.5}$$

Λ is invariant under the time evolution and it is called the *attractor* of the system. The instability caused by the stretching in the phase space may be thought of in terms of an evolution over time as an exponential divergence of nearby states when evolved by (1.4). This property is called *sensitivity to initial conditions* (SIC for the sequel) and it is the ultimate responsible agent for the apparent randomness and the implied unpredictable character of time paths picked from orbits $\{f^t(x)\}_{t \geq 0}$ of (1.4). An attractor (1.5) supporting such an unstable (chaotic) dynamics is called *strange attractor* or *chaotic attractor* [Ruelle,87]. Some classical examples of strange attractors can be seen in [Eckmann & Ruelle,85], and also in [Holden & Muhamad, 86].

We can formulate now what might be a plausible (deterministic) mechanism of generation of the signal (1.3)

The Strange Attractor Hypothesis (SAH).
The time signal (1.3) has been recorded on an unstable (that is, sensitive to initial conditions) orbit $\{f^t(x)\}_{t \geq 0} \subset \Lambda$ of a flow generated by a non-linear dynamical system (1.4) in its stationary regime, that is,

$$u_k = h(f^k(x_0)), \quad k = 1, 2, \ldots, N_p \tag{1.6}$$

where $h : \Omega \mapsto \mathbb{R}$ is some (smooth) "observable" and $x_0 \in \Lambda$ is some time-zero recording state on the attractor of the system.

1.2 Testing the SAH.

Some tests to check the SAH on the series (1.3) are built on the following considerations

1. *Aperiodic evolution.* No trivial periodicity is clear from a u_k vs. k plot, so that the signal really shows itself as random. This rules out the possibility that $\{u_k\}_k$ "lives" on a classical attractor (fixed point, limit cycle, low-dimensional torus). This aperiodic evolution should be certified by both the spectral and the autocorrelation analysis of the series. The power spectrum or power spectral density (PSD) [Priestley,81] is computed in order to detect hidden periodicities. A low–quasiperiodic regime can be discarded from this analysis, since the spectrum is concentrated in this case over a discrete set of frequencies (see figures C.4–C.6). A large number of superimposed oscillations can hardly be grasped from this analysis, and the situation worsens in a experimental context. The PSD–analysis cannot distinguish amongst a highly–excited quasiperiodic regime, a chaotic regime monitored on a strange attractor and a stochastic noise (maybe with some structure). In all those cases, an active power band of a continuum of frequencies is appreciated in the periodogram (PSD) (this is called a *broad–band spectrum*). Further work is required to address such a distinction. On the other hand, the autocorrelation analysis of a chaotic evolution must show a *fast decaying ACF.* This is because the signal quickly becomes uncorrelated, which implies a rapid loss of information about the future state evolution from the knowledge of its past.

   A visual examination of the clouds of points of the one, two, and three histories of the series (see (1.7) for a definition) often gives more information about the underlying process than many other sophisticated tests described below. Some structure can be appreciated in the embedded data when some deterministic dynamics is involved in the process evolution.

2. *Dimensionality.* The dissipative nature of the system (1.4) causes the attractor Λ to be a zero-volume subset of $\mathbb{R}^M$, which means that any considered concept of dimension of Λ yields $\dim\Lambda < M$. An estimate of (an adequate concept of) the dimension of Λ gives a measure of the dissipation in the system. Moreover, $\dim\Lambda$ gives the number of effective degrees of freedom sustaining the activity of the system in its long–term behaviour. While quasiperiodic attractors have integer dimensionality, strange attractors are usually characterized by fractional dimensions. Both phenomena can be distinguished from some stationary random processes, which have a different dimensionality behaviour.

3. *Information creation.* A significant measure of the global randomness present in the system is that concerning the rate of creation of information during the time evolution of a typical orbit of the system. This measure is represented by the concept of *entropy* (in the Kolmogorov-Sinai sense), which gives the asymptotic averaged information provided by a new observation of the state of the system [Walters,82]. From this viewpoint, purely deterministic systems are those with zero entropy whereas a positive entropy is characteristic of randomness-displaying systems, and in particular of chaotic systems.

4. *Sensitive Dependence on Initial Conditions* (SIC). This is the term used for describing the local stretching on the phase space of a chaotic system described in **1.1.2**. The (exponential) divergence of nearby orbits are well-described by the *Liapunov exponents* of the system, which are well–founded for a smooth dynamical system. A feature of chaotic dynamics is the existence of positive Liapunov exponents.

Liapunov exponents and Kolmogorov-Sinai entropy are concepts deeply associated with an underlying stationary measure for the process which is also ergodic under the time-passing transformations. Such concepts will be briefly reviewed in section **1.4**.

It should be pointed out that the properties itemized above are themselves strongly correlated. A 'true' aperiodic time path causes both a rapid loss of information, which is displayed by the ACF–analysis and by the positiveness of the entropy, and a continuous spectrum. The usual fractalness of the attractor and the production of information are both generated by the SIC–feature of the system.

The characterization just described provides the guidelines for building a testing program for chaotic dynamics in a given deterministic system when the governing equations are known. However, there still remain some important difficulties when the SAH has to be tested in an experimental contex, namely

- **A** The evolution equations driving the assumed underlying dynamical system are not usually known. The situation is even worse, since neither the dimension

of the phase space of the acting system nor the dimension of the attractor Λ are known. Therefore the problem of **reconstructring the dynamics** on Λ from the scalar time series (1.3) "observed" on the physical system must be addressed. A solution for this problem is reviewed in section **1.3** and will be applied to the sea wave signals throughout the report.

- **B** Assuming that the signal is long enough for the transients to be damped out so that the reconstructed orbit (via part **A** above) truly "lives" on Λ, properties for the whole attractor must be inferred for the observation of a single orbit. This is the problem of statistical inference, in which relevant theoretical quantities are (almost surely) computed from time averaging along a unique observed orbit. A satisfactory answer depends on the goodness of a typically registered orbit to give a faithful picture of what generically happens on the system. The tool for analyzing such a property is the **ergodic theory**, and more precisely the *ergodic theorem*, which is quoted in section **1.4**. Notice that ergodicity is also the key tool for performing spectral analysis on stochastic processes [Koopmans,74].

- **C** Some averaged-computed quantities should provide a test for discerning (at least theoretically) among *ordered* (periodic or quasiperiodic), *chaotic* (SIC–deterministic) and *purely random* systems. The quantities typically calculated are **(fractal) dimension** (chapter **2**), **Kolmogorov-Sinai entropy** (chapter **3**), and **Liapunov exponents** (chapter **4**). The tests supplied by these concepts will be presented throughout the report, as well as the results obtained when applied to our SWH–series.

- **D** Effective **numerical and computational implementation** (efficient algorithmic designing, statistical significance level, error estimates, robustness, etc.) are by no means minor problems, and should also be carefully addressed. In fact, much of the current work in the area focuses on this point. Chapters **2** to **4** are mainly devoted to this practical item.

1.3 Recovering of the Smooth Dynamics from a Scalar Signal. Embedology

In this section we address in some detail the problem of the reconstruction of a true orbit in the phase space of the underlying dynamical system from the observation of the registered scalar series (1.3). The practical problems associated with the implementation are also dicussed. In terms of the notation used to formulate the SAH (1.6), the orbit $\{f^i(x_0)\}_{i\in\mathbb{N}}$ is supposed to be recovered from the series $\{u_i\}_{i\in\mathbb{N}}$.

1.3.1 Delay Embedding Theorems.

A first idea about a successful procedure goes back to [Packard et al,1980] and was also suggested by David Ruelle. The method relies on the fact that, in a non-linear dynamical system with M degrees of freedom, *any* M independent quantities describing uniquely the states of the attractor are dynamically equivalent. This is so because a non-linear coupling among the variables causes that (almost) any observation of the system carries *all* the relevant information of the system dynamics.

The idea in [Packard et al,1980] was to choose some integer delay τ and to construct, for each $i \in \mathbb{N}$, the m-dimensional vector (called *m-history*)

$$x_i^{(m)} = (u_i, u_{i-\tau}, \ldots, u_{i-(m-1)\tau}) \in \mathbb{R}^m, \tag{1.7}$$

so that $A_m = \{x_i^{(m)}\}_i$ is a set of points embedded in $\mathbb{R}^m$. Observe that $x_i^{(m)}$ is uniquely determined for each i. Taking into account (1.6), A_m can be seen as the image of the original attractor Λ by the *delay-coordinate* mapping $J_{(f,h,m)} : \Lambda \mapsto \mathbb{R}^m$ given by

$$J_{(f,h,m)}(x) = \Big(h(x), h(f(x)), h(f^2(x)), \ldots, h(f^{m-1}(x))\Big), \tag{1.8}$$

where f^k denotes the k–fold composition $f \circ f \circ \overset{(k)}{\ldots} \circ f$. If $J_{(f,h,m)}$ is a diffeomorphism onto $J_{(f,h,m)}(\Omega)$ the systems (Λ, f) and $(A_m, J_{(f,h,m)} \circ f \circ J^{-1}_{(f,h,m)})$ are *dynamically equivalent* (i.e. differentiably conjugate), and therefore diffeomorphism–invariant properties of the unknown system are preserved in the recovered system defined on $\mathbb{R}^m$. These properties include many dimension concepts, the measure–theoretic

entropies, Liapunov exponents. Takens [Takens,81] proved in a famous theorem that *generically* $J_{(f,h,m)}$ is actually a diffeomorphism provided that $m \geq 2M+1$. A generic property on a set X is that which holds for a dense and open subset of X.

> **Takens's theorem** (1980). Let Ω be a compact manifold of dimension M. For pairs (h, f), where $f : \Omega \mapsto \Omega$ is a smooth (at least C^2) diffeomorphism and $h : \Omega \mapsto \mathbb{R}$ is a smooth observable, it is a *generic* property that the mapping $J_{(f,h,m)} : \Omega \mapsto \mathbb{R}^m$ is an *embedding* (i.e. $J_{(f,h,m)}$ is a diffeomorphism onto its image, which is besides a submanifold of $\mathbb{R}^m$) provided that $m \geq 2M+1$.

There is a version of the above theorem for continuous-time flows. There is also a version for non-compact Ω provided that the observable h is a proper function.

Notice that Takens's theorem is not directly applicable to strange attractors since they typically are not smooth manifolds, but fractal sets. This poses the problem of finding a (minimal) manifold containing the attractor, and finding its dimension in order to be sure about the dimension that will suffice for a faithful reconstruction of the attractor and the dynamics taking place on it. Unfortunately if a, say 2.4–dimensional, attractor of a, say 10–dimensional flow, cannot be guaranteed to be contained in a smooth manifold of dimension less than 10, Takens's theorem only asserts that a 'good' reconstruction is available from $\mathbb{R}^{21}$ up!. The problem of course worsens when Λ is a low–dimensional attractor of an infinite–dimensional flow. Even when, via dissipation, a suitable finite (but high) dimensional manifold containing Λ can be found, the minimum dimension (supplied by Takens theorem) for $J_{(f,h,m)}$ to be an embedding could be prohibitive. Moreover, there are no criteria for finding the smallest manifold containig A.

A second problem about the applicability of Takens-type theorems concerns *genericity*. This is because generic sets can be negligible in terms of probability, and thus of improbable occurrence. There are examples from dynamical system theory (see [Sauer et al,91]) of generic sets with arbitrarily small probability.

Both an extension of Takens theorem to non-smooth compact sets embedding and

a Takens-type theorem with the assertion "generically" in the statement of Takens theorem replaced with "almost surely" (in a suitable sense) have recently been put through by Sauer, Yorke and Casdagli in their excellent article [Sauer et al,91]. They also extended previous celebrated results of Whitney on manifolds embedding and Mañé on injective projections of manifolds. We cite below what they called the *fractal delay embedding prevalence theorem for diffeomorphisms*, which generalizes the theorem of Takens in the two senses commented above. We denote by $\dim_{BOX}(\cdot)$ the box-counting dimension, which is defined in **2.1.2.**

> **Theorem. (Fractal Delay Embedding Prevalence Theorem** (1991)) Let f be a diffeomorphism on an open set $U \subseteq \mathbb{R}^M$, and let A be a compact subset of U such that $\dim_{BOX}(A) = d$. Let $m > 2d$ an integer. Assume that for every integer $p \leq m$, the set A_p of periodic points of A of period p verifies $\dim_{BOX}(A_p) < p/2$, and that the linearization Df^p for each p-periodic orbit has distinct eigenvalues.
>
> Then for "almost every" smooth (C^1) observable $h : U \mapsto \mathbb{R}$, the delay coordinate map $J_{(f,h,m)} : U \mapsto \mathbb{R}^m$ defined in (1.8) is:
>
> i) Injective on A.
>
> ii) An embedding of each compact subset of a smooth manifold contained in A.

Some comments are in order:

a) The statement "almost every" in the theorem must be understood in the sense of *prevalence*. This is a concept designed by some of the authors in [Sauer et al,91] in order to traslate the concept of "with probability one" to infinite–dimensional spaces (as the space of smooth maps or delay–coordinate maps are). For subsets of finite–dimensional spaces the term *prevalent* means "almost every" in the sense of 'outside a set of zero measure'. Roughly, a Borel subset P of a Banach space is *prevalent* if any element of the space perturbed by an element of a suitable finite-dimensional subspace E lies in P for almost every element of E (with respect to the Lebesgue measure in E). Prevalence in general implies denseness. The strengthening of prevalence with respect to genericity is clear: not only are there embeddings near

a given smooth map, but *almost every* (in the sense of prevalence) map near such a smooth map is an embedding.

b) Note that the theorem allows us to map a d–dimensional attractor without self–intersections into $\mathbb{R}^m$, for $m > 2d$, independently of the dimension of the flow whose asymptotic dynamics generated the attractor. Thus a 2.4–dimensional attractor is injectively mapped into $\mathbb{R}^5$ with probability one, and also any differential structure of a manifold contained in the attractor is also preserved by almost every smooth mapping to $\mathbb{R}^5$.

c) On the other hand, the restrictive (but mild) conditions on the diffeomorphism f cannot be relaxed. In [Sauer et al,91] there are examples of periodic attractors which cannot be injectively mapped into $\mathbb{R}^m$ by any delay–coordinate map $J_{(f,h,m)}$. Observe that the assumptions on the number of periodic points allow a finite number of equilibria and p–periodic orbits (though not countable in general). The extra hypotheses are a consequence of the fact that the class of delay–coordinate maps is not wide enough to perturb non–embedding maps in order to get (generically) embedding maps still within the class. Notice that there are no restrictions when the wider class of smooth maps is considered (see extensions of Whitney theorem in [Sauer et al,91]).

d) Notice that the theorem is also valid for continuous flows ϕ_t by taking $f \simeq \phi_\tau$ for a chosen $\tau > 0$. Under smooth conditions, the map f is a diffeomorphism and the theorem applies for every delay time τ which respects the hypothesis on the periodic orbits in the theorem. In general, a single orbit displaying a strange attractor (under hypothesis of ergodicity) does not contain (low-period) periodic orbits at all, so that it can be mapped injectively into $\mathbb{R}^m$ as soon as $m > 2\dim_{BOX}(A)$.

Some comments regarding practical applications of embedding theorems are in order. The following section collects some of them.

1.3.2 Implementation: Embedology.

Of course, in the experimental setting neither the actual phase space dimension dimΩ nor the dimension of the attractor are known *a priori*, and so an m-dimensional embedding of the series should be tried for each m. Let the pair (f, h) be fixed. For each positive integer m, let J_m denote the mapping

$$J_m : \Omega \mapsto J_m(\Omega) \subset \mathbb{R}^m$$

as defined in (1.8), and let A_m denote the reconstructed attractor $J_m(\Lambda)$. Let $\hat{f}_m \stackrel{\text{def}}{=} J_m \circ f \circ J_m^{-1}$ and call the pair

$$(A_m, \hat{f}_m) \tag{1.9}$$

the *m-dimensional (Ruelle-Takens) reconstruction* (or *m-dimensional embedding*) of the series (1.3). Takens's theorem asserts that a faithful conjugate picture of the original compact phase space is typically rendered by the m-embedded set A_m as soon as $m \geq 2\text{dim}\Omega + 1$, and the fractal delay embedding theorem ensures that a one-to-one image of the attractor Λ is provided by A_m from 2dimΛ up. This is the reason why it has been widely recognized in the literature [Eckmann & Ruelle,85] that a picture of the attractor is quite fair for reconstruction values of m below $2\text{dim}\Omega + 1$. An easy example is provided by equation (1.2), where a 2-dimensional delay reconstruction reveals the non-linear dependence in a series from (1.2) which passed statistical tests for independence. Of course, the phase-space structure is not lost in a higher dimensional embedding (see figures C.7–C.9).

On the other hand, the choice of a reconstruction dimension m to study the set $J_m(A)$ (the dynamics f and the observable h are fixed) depends on the issue being investigated. In chapter **2** we will be concerned with dimension computation. If a metric dimension (see **2.1.2** below) is what is intended to be computed, all that is needed is

$$\text{dimension } (J_m(A)) = \text{dimension } (A). \tag{1.10}$$

There is a result by Pertti Mattila [Mattila,75] asserting that equality (1.10) holds for almost every orthogonal projection on $\mathbb{R}^m$ provided that $m \geq \text{dim}A$, where dim$(\cdot)$ stands for Hausdorff dimension (see **2.1.2**). It is shown in [Hunt et al] that

(1.10) actually holds for almost every measurable map (in the sense of prevalence) as long as $m \geq \dim A$.

To compute the subtler dimension concepts associated with the invariant measure (see **1.4**) a measure spaces isomorphism is necessary, and thus J_m must be injective. The theorem above ensures that this occurs for $m \geq 2\dim_{BOX}(A) + 1$. We will address dimension concepts and dimension computation in chapter **2**.

Since all the differential information on the manifold Ω must be preserved in order to compute Liapunov exponents, Takens's theorem permits this computation from the cloud of m–histories embedded in $\mathbb{R}^m$ provided that m> $2\dim\Omega$. See chapter **4** for definitions and a precise statements of these facts.

The property that, generically, dynamical (or geometrical) quantities of the true attractor are preserved by a suitable Ruelle-Takens reconstruction has been termed *Invariance Principle* in [Brock,86]. In words of [Takens,81] the metric properties of the closure of a true orbit, $\mathrm{cl}(\{f^i(x_0)\}_{i\in\mathbb{N}}) \subset \Lambda$, are the same (in the sense that the ratio between distances in Ω and the image distances in $\mathbb{R}^m$ is uniformly bounded above and below) as those of the m–embedded sequence $\mathrm{cl}(\{f^i(x_0)\}_{i\in\mathbb{N}}) \subset A_m$.

For practical purposes two parameters have to be chosen to construct (1.9), namely, the delay time τ (an integer value in our setting) and the reconstruction dimension m. As commented above, nothing accurate can be said about the choice of m, so we must reconstruct A_m for some small values of m and look for some structure in the phase plot or some *ad hoc* projection. Even though the construction (1.7) should give a good picture of Λ for a generic τ, some care has to be taken for practical choices. Some criteria have been addressed so far ([Fraser & Swinney,86], [Buzug et al,90], [Liebert et al,91] or [Liebert & Schuster,89]) for obtaining optimal delay reconstructions. A different idea for dynamics recovering from scalar data based on Singular Value Decomposition (SVD) was addressed in [Broomhead & King, 86]. A comparison study [Fraser,89] between the SVD approach and the criterium of mutual information proposed in [Fraser & Swinney,86] showed the latter one to be superior. In [Grassberger et al,91], [Theiler,90] it is advised to choose (τ, m) such that the product $(m-1)\tau$ is slightly below a typical decorrelation time, i.e.

$$(m-1)\tau \simeq T_{dec} \tag{1.11}$$

where T_{dec} is the first zero of the ACF of the signal. This criterium makes sense because a small $\tau(m-1)$ causes the m-histories x_i and x_{i+1} to be highly correlated and the data set A_m is spuriously stretched along a hyperplane (this fact has been widely documented, e.g. [Theiler,90], [Grassberger et al,91]). On the other hand, a choice of (m, τ) such that the product $(m-1)\tau$ is too large will give an unfair, random-looking reconstruction due to uncorrelation effects.

Our approach will be essentially that of [Grassberger et al,91], where it is recommended to take $\tau = 1$ fixed, to increase the embedding dimension m and to make the calculations of the testing invariants at each m considering spatial–correlation only from a dynamical (and not temporal) point of view, as suggested in [Theiler,90]. This simply means that points that are neighbors just because they are close in time should not be considered when calculating any statistical quantity. We will return to this issue in section **2.2.2.**

Takens's theorem shows that if the series $\{g(u_i)\}_i$ is considered, where g is some generic smooth transformation, estimates of the invariants from a m–embedded data set obtained from this new series must give the same result as those obtained from the original data set $\{u_i\}_i$. In practice, g can be taken to be some specific transformation relevant to the problem under consideration. This yields a robustness test for the estimates to be obtained from the signal. In [Lorenz,90] it is shown that certain estimates vary significantly depending on the coupling among the variables on the system, and also on what variable or observable is monitored to reconstruct the dynamics. This is also a reason to perform some robustness test based on genericity. See section **A.2.2.**

Several well–known dynamical systems (Lorenz and Rössler 3-dimensional flows and Henon mapping) as well as white noise (gaussian or uniform) processes will constitute a testing laboratory for the computational work we have developed as an implementation of the theoretical ideas. The phase portraits of these systems are given in section **C.1** and the m–histories obtained from a scalar time series monitored in their time evolution can be seen in section **C.4.**

1.4 Ergodic Theory. Statistical Analysis on Dynamical Systems.

Ergodic theory is the tool for performing statistical analysis on dynamical systems. The situation is quite similar to that of statistical inference in stochastic processes, where some general theoretical properties are obtained from measurements on a single (long enough) realization of the acting process. Ergodic theory can be applied indeed to stochastic processes. The combination of dissipation and SIC in the SAH context (see section **1.2**) makes the (*a priori* more natural) geometrical or topological description of the attractor extremely difficult. A more adequate approach is that of measure theory, where the typical behaviour of the system can be grasped.

The object of study of ergodic theory is the dynamics on probability spaces. The sample space in this probabilistic scheme could be either the phase space Ω of the dynamical system (1.4) (deterministic case) or the probability space of an stochastic process under the dynamics of the time flow. In the deterministic case, transients are irrelevant, so that the attractor Λ is the (long–term) sample space. Now a probability measure must be defined on Λ. It could be considered the Lebesgue M-dimensional measure ($L^M(\cdot)$), but the dissipativeness of the system gives $L^M(\Lambda) = 0$, which makes it useless to develop such a probabilistic analysis on Λ. The useful concept for such an analysis is that of *stationarity*, which is read in terms of a probability measure by means of time–invariance.

Time invariance. A (probability) measure μ is *invariant* (or *stationary*) for the evolution semigroup $\{f^t\}_{t\geq 0}$ if

$$\mu(f^{-t}(E)) = \mu(E), \tag{1.12}$$

for $t \geq 0$ and E μ–measurable.

The first question that must be solved is that of the existence of invariant measures for a discrete system (Λ, f). When Λ is compact and f is continuous this is a well-known result from ergodic theory (see e.g. chapter six from [Walters,82]). Observe that no assumption on the differentiable structure of Λ is needed for these results

of measure-theoretical type. For readers more familiar with the theory of stochastic processes μ can be read as the stationary measure of the process $\{f^t\}_t$ and Λ as its support. See section **4.4**.

The next goal is to assure that time series analysis can be performed on a single orbit, i.e. theoretic-probabilistic quantities can be inferred from *time averages* computed along an orbit/realization of the system. This is granted when the system is *indecomposible* or *ergodic*, that is, all time–invariant sets are measure–theoretically trivial (i.e. are μ-full or μ-null sets). A set E is said to be *invariant* if $f^{-t}(E) = E$ for all $t \geq 0$.

> **Ergodicity.** (μ, f^k) is *ergodic* (or *indecomposable*) if there are not invariant sets E_1 and E_2 with positive non–full μ-measure, i.e. E_i such that $f^{-k}(E_i) = E_i$ fot all $k \geq 0$ and $0 < \mu(E_i) < 1$, $i = 1, 2$.

A differentiable dynamical system carries at least one, and possibly many, ergodic measures. We provide an example below, of a dynamical system carrying a continuum of invariant ergodic measures.

In the above context, the following celebrated theorem (see e.g. [Walters,82] or [Billingsley,78] for a proof) allows statistical analysis to be performed on Ω

> **Birkhoff's ergodic theorem** (1927). Let μ be an invariant measure for the discrete system (Λ, f), and let h be a μ–integrable observable function, then the set of points where the limit
>
> $$\lim_{k \to +\infty} \frac{1}{k} \sum_{j=1}^{k} h(f^j(x)) = h^*(x) \tag{1.13}$$
>
> exists has full μ–measure, and furthermore
>
> $$\int_\Lambda h d\mu = \int_\Lambda h^* d\mu.$$
>
> Moreover, if (Λ, f) is ergodic, then $h^*(x) \equiv$ *constant* on a set of full μ–measure, and therefore
>
> $$\lim_{k \to +\infty} \frac{1}{k} \sum_{j=1}^{k} h(f^j(x)) = \int_\Lambda h d\mu \tag{1.14}$$

for a set of full μ–measure.

The ergodic theorem shows that an (invariant) ergodic measure in Λ can be defined by time averages, i.e. the measure of a set E can be calculated by

$$\begin{aligned} \mu(E) &= \lim_{k\to\infty} k^{-1} \sum_{j=1}^{k} 1_E(f^j(x)) = \\ &= \lim_{k\to\infty} k^{-1} \sharp\{j:\ f^j(x) \in E,\ 1 \le j \le k\}, \end{aligned} \tag{1.15}$$

where $1_E(\cdot)$ stands for the characteristic function of the set E and x belongs to a set of full μ-measure. As remarked above a strange attractor typically carries many different ergodic measures, so some criterion should be used to select one of them. As an example, the invariant set $I = [0,1]$ (observe that this is not formally an attractor) associated with the saw–tooth dynamics (1.2) supports uncountably many ergodic measures: Identify the unit interval with the product space $\{0,1\}^{\mathbb{N}}$ (this can be done biunivocally except for a countable set). Let μ_p be the Bernouilli (product) measure on I associated with the measure $\mathbf{p}$ on $\{0,1\}$ defined by $\mathbf{p}(0) = p$ and $\mathbf{p}(1) = 1 - p$, $0 \le p \le 1/2$, that is

$$\mu_p = \mathbf{p}^{\mathbb{N}} = \mathbf{p} \times \mathbf{p} \times \ldots. \tag{1.16}$$

It can be seen [Walters,82] that μ_p is invariant and ergodic for the mapping defined by (1.2). As a suggestion for choosing one of such measures, observe that the unique choice of p which makes μ_p absolutely continuous with respect to the Lebesgue measure is $p = 1/2$ ($\mu_{1/2}$ is in fact the Lebesgue measure). From a physical viewpoint this is the relevant measure.

The next step towards a chaos testing analysis of the signal (1.3) is that issued in the item **C** of section **1.2**) and is two–folded. First, a theoretical testing concept to be defined must be invariant under conjugation, since the analysis will be performed on the reconstructed system (1.9) and the invariance principle in section **1.3** must be applied. On the other hand, such invariants should have a detectable distinct behaviour when evaluated on random systems and when they are evaluated on deterministic low-dimensional systems. As explained above, the rest of this report will be devoted to the design, theoretical foundation, and practical implementation of several of those statistical invariants.

Chapter 2

Testing Dimension. The Correlation Dimension

In this chapter we firstly review the concept of fractal dimension from the point of view of the role it plays in the SAH context. The usefulness of the different dimension concepts to characterize both the fractal geometry and the dynamic features of strange attractors is then reviewed, as well as their practical implementation. A second section focuses on the computatibility and estimate aspects of the widely used *correlation dimension*, introduced by Grassberger and Procaccia [Grassberger & Proccacia,83a]. The mechanism of the box–assisted technique designed by Grassberger for the efficient computation of the correlation dimension is thoroughly unraveled. The method used to obtain Significant Wave Height (SWH) time–series, along with the results of the dimension test on SWH–series are reported in the third section of this chapter. In appendix **A** we review further test–designing from concepts associated with the invariant measure of the system (see chapter **1**), that can be computed from the algorithm developed by Grassberger.

2.1 Dimension of (Strange) Attractors.

2.1.1 Characterization of Attractors by Dimensionality.

A primary knowledge to characterize the attractor's nature is that of its *dimensionality* in some sense. We give below some arguments supporting such a statement. For a more thorough study on attractor dimension we suggest [Farmer et al,83], and [Theiler,90] for a more general approach to the concept of fractal dimension. For the discussion, it might help to have in mind that any reasonable concept of dimension applied to the classical attractors; fixed point, limit cycle or p-torus will give the expected values, that is 0, 1 and p respectively.

- *Modelling variables.* To begin with, the dimension gives the minimum number of essential variables necessary to model the (asymptotic) dynamics of the system. In other words, the dimension of an attractor represents *grosso modo* the information necessary to locate, within a given accuracy, the position of a state on the attractor. Therefore, an estimate of the attractor's dimension will give at least a lower bound for the number of dynamically relevant modelling variables. Of course, $\dim\Lambda \leq M = \dim\Omega$ but this certainly does not give relevant information.

- *Dissipation.* Dimension also measures the degree of dissipation which takes place in the system. As a consequence of the dissipation (to be read in terms of phase space volume compression) the attractor has zero M-dimensional volume and the dimension of the attractor gauges the long–term effect of the system compression.

- *Geometric and dynamical strangeness.* When a chaotic dynamics operates on the phase space, the stretching–plus–folding mechanism (see **1.1.2**) provokes the attractor to be a very intricate geometric structure, a sort of Cantor–like set (far from being a smooth manifold), when the mechanism is developed *ad infinitum.* It can be said that a strange dynamics causes a strange geometry, what is usually termed a *fractal* geometric structure. Fractal Geometry is the field which deals with geometric and measure–theoretic properties of irregular,

highly non-smooth sets like those attractors associated with a chaotic dynamics; and so it provides the natural tools for characterizing chaotic attractors both globally and locally. It is worth observing that dynamical and geometric strangeness are not necessarily related to each other (see some examples in [Eckmann & Ruelle,85]), even though the latter is usually a consequence of the former in the context of the SAH. The dimension of a fractal attractor is typically a non-integer value, even though it can be an integer figure in some theoretically–designed examples.

2.1.2 Different Types of Dimension.

There are some different manners of dimensioning an attractor, which arise from metric, probabilistic or dynamical considerations. These are introduced in the rest of this chapter.

Metric dimension.

Such dimension concepts are related to the geometric strangeness mentioned above and are thus only concerned with the way in which the attractor set occupies its geometric ambient space. No consideration from the dynamic genesis of the set is taken into account. Metric dimensions can be computed both from global and from local techniques.

Globally, the metric dimension of a set E is found as a critical value in the power–like behaviour of efficient coverings of E . Specifically, fix some resolution scale $\delta > 0$ and find the infimum of the sums

$$\sum_{\substack{U \in \mathcal{U} \\ \mathcal{U} \in \mathcal{C}(E,\delta)}} |U|^a, \tag{2.1}$$

for each $a > 0$, where $|\cdot|$ stands for the diameter of a set, and $\mathcal{C}(E,\delta)$ is the collection of countable coverings of E by some class of sets U (open balls, dyadic

intervals, etc.) with a controlled size, typically

$$| U |= \delta. \tag{2.2}$$

Let δ go to zero and find the unique positive d such that (2.1) is infinite for $a < d$ and is zero for $a > d$. Such a value d is called the *box-counting* or *Minkowski dimension* ($\dim_{\mathrm{BOX}}$) of E, and it is the most extensively used for dimension estimates in applied sciences because of its easy computability.

The local approach to metric dimension is related with the existence of a Borel measure ν such that $\nu(E) > 0$ and whose *logarithmic density* at each point x (also called *pointwise dimension*),

$$\alpha(x) \equiv \alpha_\nu(x) = \lim_{r \to 0} \frac{\log \nu(B_r(x))}{\log r} \tag{2.3}$$

exhibits a uniform behaviour on E (see e.g [Falconer,90]). Observe that the density in (2.3) gives the scaling behavior of the measure ν on small balls $B_r(x)$, that is, for r arbitrarily small

$$\nu(B_r(x)) \sim r^{\alpha(x)}. \tag{2.4}$$

Commonly, there is a 'natural' measure supported by E whose local behaviour (equation (2.3)) on E can be understood and so dimension estimates can be drawn. The metric dimension related to the behaviour of (2.3), however, is not the box–counting dimension, but the more involved concept of *Hausdorff dimension.* Hausdorff dimension is obtained as a threshold value in (2.1), but the infimum is evaluated on a wider class of sets. More precisely, sets U with $| U |\leq \delta$ (instead of (2.2)) are allowed in (2.1), and therefore Hausdorff dimension might be smaller than box–counting dimension. The lower logarithmic density and Hausdorff dimension are intimately connected for a wide class of sets (see e.g. [Falconer,90], [Tricot,82]), but if $\alpha(x) = \alpha$ uniformly on E in (2.4) it can be proved [Tricot,82] that box–counting and Hausdorff dimension coincide. When this equality cannot be assured, the major drawback of box-counting dimension must be taken into account, namely it assigns the same value to a set A and to its closure cl$A$,

$$\dim_{\mathrm{BOX}}(A) = \dim_{\mathrm{BOX}}(\mathrm{cl}(A)). \tag{2.5}$$

For a rigorous treatment on both theoretical and applied aspects of a number of concepts of (metric) dimension see [Falconer,90].

Measure dimension.

Even though the metric dimension of a strange attractor can elucidate relevant geometric aspects of the dynamics, it may not tell much about the dynamics itself. The attention must be turned to the stationary measure that carries the information about the dynamics. A concept of dimension dynamically relevant must involve the probabilistic nature of the system. There is another reason to focus on the invariant measure of the system: in order to compare the dimensionality behaviour of deterministic processes with that of stochastic processes a concept of dimension on stochastic systems should make sense. This concept naturally arises associated with the stationary measure of the process (see section **4.4.1**).

Dimension of the measure μ. It is defined as the infimum of the Hausdorff dimensions of sets displaying μ significantly, that is

$$\dim\mu = \inf\{\dim A : \ \mu(A) > 0\}, \tag{2.6}$$

where $\dim(\cdot)$ stands for Hausdorff measure. It can be seen that the same value is attained in (2.6) over the smaller collection of sets of full μ-measure provided that μ is ergodic.

The dimension of the measure μ gives the size of a 'core' set sustaining the dynamics of the system (think of μ as obtained by averaging (see (1.15)), whereas the box–counting dimension works on the support of the measure, supp μ, which can be dynamically uninformative. An illustrative example is provided by the ergodic pairs $([0,1],\mu_p)$ associated with the saw–tooth mapping F in equation (1.2) (see section **1.4**). It is known [Deliu et al,91] that

$$\dim\mu_p = -\log_2(p^p(1-p)^{1-p}) \tag{2.7}$$

for $p \in (0,1/2)$, whereas dim(suppμ)=dim $[0,1]$=1. Moreover, the box–dimension of an orbit $O(x) = \{F^i(x)\}_i$ observed in the system $([0,1],F,\mu_p)$ would yield $\dim_{\mathrm{BOX}}(O(x)) = \dim_{\mathrm{BOX}}(\mathrm{cl}(O(x))) = 1$ with μ_p–probability one independently of the underlying probability μ_p. This is so because of the property (2.5) of the box dimension and the fact that, for each p, μ_p-almost any orbit is dense in the interval. It is possible to choose values of p so that the 'true' dynamics (with respect

to the measure μ_p) takes place on a set of Hausdorff dimension arbitrarily close to 0, what is certainly reflected by the measure dimension analysis in the formula (2.7).

The following theorem [Young,82] makes a connection between the uniform behaviour of the logarithmic density $\alpha_\mu(\cdot)$ and the dimension of μ. More interestingly, combined with the ergodic theorem provides a practical tool to calculate the dimension of the invariant measure computed by averaging (see (1.15)).

Young's theorem (1982). Let μ be a probability measure supported in A, and let $\alpha \geq 0$ such that

$$\alpha_\mu(x) = \alpha \tag{2.8}$$

for a set of full μ-measure, then $\dim\mu = \alpha$.

In fact, the hypothesis of the theorem suffices to show that, for some cases, many other 'probabilistic' dimensions take also the value α [Young,82].

The existence of the limit $\alpha_\mu(x)$ in (2.8) for μ-almost any x does not imply, in general, that this limit has to be constant μ-almost everywhere [Cutler,90a]. However, ergodicity makes things simpler for the case of an invariant measure on a smooth dynamical system.

Cutler's theorem (1990). Let (Ω, f^t, μ) be an ergodic dynamical system. If the logarithmic density of μ, $\alpha_\mu(\cdot)$, exists on a set of full measure, then it is constant μ-almost everywhere.

Cutler's theorem supplies a test of ergodicity for a signal recorded on a dynamical system by computing the logarithmic density distribution on Λ (assumed to exists on a set of full measure). This involves the calculation of statistical quantities as the sample average $\overline{\alpha_\mu} = < \alpha_\mu(x) >_{x\in\Lambda}$, or the sample variance over an ensemble or the whole data set. A 'gaussian–like' appearance of the histogram of the logarithmic

density enhances the hypothesis that the system is ergodic. This hypothesis, however, can be very difficult to check in practice. On the other hand, a non–gaussian appearance of the histogram may be due to a richer behaviour of the mass-spreading of the measure μ, thus suggesting the presence of two or more dimensional ergodic components mixed in the system. This is a suggestion of non–uniformity in the distribution of μ. The measures displaying such behaviour are called *multifractal measures* [Halsey et al,86]. A branch of the non-linear analysis of time series, called *multifractal analysis*, studies this kind of measure.

Of course, the analysis above must be developed on some Ruelle–Takens reconstruction, and thus the measure structure of the original system must be preserved by the mapping $J_{(f,h,m)}$ in (1.8) provided that m is large enough. It can be proved that, under the SAH, the logarithmic density distribution of the measure μ is preserved when computed from the available 'reconstructed' measure $\hat{\mu}_m = \mu \circ J^{-1}_{(f,h,m)}$ (see **1.3** for notation).

Liapunov dimension.

This is a concept of dimension, introduced in [Kaplan & Yorke,78] as a lower bound for a metric dimension, intrinsically associated with the dynamics of the system. It is calculated from the Liapunov exponents of the system, which measure the deformation of the tangent bundle under the time evolution. The Kaplan-Yorke conjecture [Frederickson et al,83], well supported by experimental evidence, claims the coincidence between the Liapunov dimension and the measure dimension for a 'typical' chaotic attractor. This dimension has been employed to detect coupling between the variables in a system [Lorenz,90]. We will discuss this dynamical dimension in more depth in chapter **4**, where Liapunov exponents will be reviewed in detail.

2.1.3 Dimension Test for 'Chaos vs. Randomness'.

The first well–known dimension test devised for distinguishing randomness from chaos in time series [Grassberger & Procaccia,83c] is that based on the fact that

m-histories (delay mappings are not embeddings in this case) of purely (stationary) random processes geometrically fill the ambient space $\mathbb{R}^m$ where they are contained, and thus any concept of dimension applied to the m-histories data set will yield the dimension m of the whole space. On the contrary, a saturation on the value of the dimension of data set in a plot of 'estimated dimension' vs. 'embedding dimension' might suggest the existence of chaotic dynamics. This test on dimensionality will be applied in this chapter and chapter **3** to sea swell time series.

The test can be summarized as follows. Let $\dim(m)$ denote the dimension (one of them) of the m-dimensional delay reconstruction of the series (1.3), then

$$\begin{aligned} (Pure)\ Random\ System \qquad & \dim(m) = m \quad for\ all\ m \\ Chaotic\ System \quad & \dim(m) = \begin{cases} m & \mathrm{m} < \dim\Lambda \\ \dim\Lambda & \mathrm{m} \geq \dim\Lambda \end{cases} \end{aligned}$$

where Λ is the strange attractor of the chaotic system. Some care has to be taken in order to apply the dimension test above, since some stochastic (stationary and non-stationary) processes are known to yield a convergent value of $\dim(m)$ [Osborne & Provenzale,89]. Conditions on stochastic process for successively filling the ambient space have been pursued [Brock,86]. From our research it also follows that much care has to be taken to consider the dimension test by itself as a basis for accepting the SAH. We show in chapter **3** that some pure random processes pass the dimension test as low–dimensional dynamical systems.

2.2 Procedures for Measure Dimension Computation.

The main theme in this section is the correlation dimension estimate problem as computed by the Grassberger–Procaccia algorithm. We firstly remark on the computational aspects of estimating box–dimension.

The problem of computability of box-counting dimension will not be addressed in this work since it has been recognized for some time [Russell et al,80], [Froehling et al,81], [Greenside et al,82] that severe problems arise in its numerical estimates when the

dimension is not very low. This comes from the exponentially increasing need of data points as the reconstruction dimension increases. We developed an efficient code, which incorporates the box–assisted mechanism [Mera et al,93] for neighbour search (see **2.2.4**), for the calculation of box–counting dimension. Our program accurately computes the box–dimension of self–similar sets (which is known to coincide with the Hausdorff dimension) and some low–dimensional attractors, but it has serious difficulties to render the dimension of space–filling sets (sets with the dimension of the ambient space) even in dimension two. Moreover, the data requirements for a significant estimate of sets with dimension greater than two are easily seen to become excessive for field registered signals.

2.2.1 Testing Correlation Dimension. The Grassberger-Procaccia Algorithm.

The correlation dimension was originally introduced in [Grassberger & Proccacia,83a]. Its success is partly due to the fact that the correlation dimension of a presumed attractor can be estimated up to a high embedding dimension, e.g. of twenty, without too much computational effort. This should be compared with the lower efficiency of an algorithm for the computation of the box–dimension estimating (see above). On the other hand, a drawback of the correlation dimension is that there are important facts in its theoretical foundation which remain unsolved.

Let m be an integer, and let $A_m = \{x_1^{(m)}, x_2^{(m)}, \ldots, x_N^{(m)}\} \subset \mathbb{R}^m$ be the finite data set on a m-dimensional delay reconstruction of the scalar signal

$$\{u_i\}_{i=1,2,\ldots,N_p} \tag{2.9}$$

(see chapter **1**). Note that $N \equiv N(m) = N_p - (m-1)$. For convenience, we will also denote the embedded set $J_m(\Lambda)$ by A_m. Young and Cutler's theorems from chapter **1** suggest estimating the logarithmic density (also called *pointwise dimension*) $\alpha_\mu(\cdot)$ of the invariant measure μ, defined in (2.3). Note that we are in fact considering on the system $(A_m, \hat{f}_m)$ the induced measure $\hat{\mu}_m = \mu \circ J_m^{-1}$ (see section **1.3**), but it can be proved that generically $\alpha_{\hat{\mu}_m} = \alpha_\mu$ $\mu - a.e.$ provided that m is large enough. We will write μ to denote any of both equivalent measures.

We assume that the logarithmic density exists μ-a.e. and moreover

$$\alpha_\mu(x) = \alpha \text{ for almost all } x \in A_m. \tag{2.10}$$

Then μ is called *exact-dimensional* [Cutler,90a] and α is called the *information dimension* of μ. Of course $\dim\mu = \alpha$ from Young's theorem, and ergodic theory allows us to estimate $\dim\mu$ as follows.

Consider, for $r > 0$

$$c_i^m(r,N) = \frac{1}{N}\sum_{j=1}^{N} \Theta(r - \| x_i^{(m)} - x_j^{(m)} \|) = \tag{2.11}$$

$$= \frac{1}{N}\sharp\{j : 1 \le j \le N,\ \| x_i^{(m)} - x_j^{(m)} \| \le r\},$$

where Θ denotes the Heaviside function and $\|\cdot\|$ stands for the maximum norm in $\mathbb{R}^m$. From the ergodic theorem it follows that μ–a.e.

$$c_i^m(r) \stackrel{\text{def}}{=} \lim_{N\to+\infty} c_i^m(r,N) = \mu(B_r(x_i^{(m)})), \tag{2.12}$$

and from (2.10)

$$\lim_{r\to 0} \frac{\log c_i^m(r)}{\log r} \stackrel{\text{def}}{=} \alpha_i(m) \equiv \dim\mu, \ \ \mu - a.e. \tag{2.13}$$

Equation (2.13) is the basis for an algorithm computing the distribution of the logarithmic density on the support of μ, which opens the door towards a multifractal analysis of the measure (see [Halsey et al,86]). There are, however, more efficient methods for estimating the multifractal spectrum that stems from a basic code computing correlation dimension [Grassberger,83].

The correlation integral $C^m(r)$ is defined by averaging in (2.11) before looking for a scaling behaviour as $r \to 0$, so that $C^m(r)$ reflects a global property of the measure. Specifically, we define

$$C^m(r,N) = \frac{1}{N^2}\sum_{i,j=1}^{N} \Theta(r - \| x_i^{(m)} - x_j^{(m)} \|) \tag{2.14}$$

and then the *spatial correlation integral* is given by

$$C^m(r) = \lim_{N\to+\infty} C^m(r,N). \tag{2.15}$$

Notice that the statistic $C^m(r, N)$ (also called the *Grassberger–Procaccia estimator*) approximates the average of the μ–mass of a ball of radius r from a sample of N data points.

The *correlation dimension* $\nu_m(\mu)$ of the measure μ (calculated from the data set A_m) is the power–like scaling law of the m–dimensional correlation integral, i.e.

$$\nu_m(\mu) \equiv \nu(\hat{\mu}_m) = \lim_{r \to 0} \frac{\log C^m(r)}{\log r} \tag{2.16}$$

provided that the limit exists.

It can be shown that the correlation dimension in (2.16) takes the same value when evaluated from different equivalent norms of $\mathbb{R}^m$. [Brock,86].

[Grassberger & Procaccia,83c] calculated the exponent ν for a number of finite and infinite well–known dynamical systems. They found heuristically that

$$\nu \leq \alpha \tag{2.17}$$

where α and ν are the information and correlation dimension of the measure μ respectively, and m is taken large enough. They also conjectured that typically an equality holds in a smooth ergodic dynamical system. [Cutler,91] has proved that (2.17) in fact holds for a Borel measure on a compact set.

On the other hand, there are examples [Cutler,90] of smooth ergodic dynamical systems in $\mathbb{R}^M$ such that $\sigma = M$ while ν is arbitrarily close to zero. Thus there is no relationship between ν and the dimension of the smallest sets displaying μ.

In order to enhance the role of the correlation dimension as a discerning tool for chaos vs. stochasticity it would be important to know what kind of stochastic processes do behave as 'space–filling' from the correlation dimension viewpoint. We say that a process is space–filling if it has full correlation dimension for all dimension m where it is embedded. In this sense it is worth mentioning that in [Brock,86] some conditions of 'regularity' for a dynamical system with noisy observer and/or noisy law of evolution were given to ensure that $\nu_m(\mu) = m$ for all m almost surely. To disentangle some noise effects corrupting deterministic systems some considerations

about the structure of the correlation integral are useful. See e.g. "the knee effect" below, which seems to have been first documented in [Ben-Mizrachi et al,84].

It was first believed that the family $\{\nu_m(\mu)\}_m$ had a different behaviour on deterministic and stochastic series. [Osborne & Provenzale,89] showed that certain correlated stochastic processes can yield a finite estimate of the correlation dimension. This is a remarkable fact [Theiler,91], that should be taken into account when interpreting correlation dimension estimates. Some items of care to be taken when applying the correlation dimension test to non–stationary signals (as those originally proposed in [Osborne & Provenzale,89]) were then proposed [Grassberger et al,91], [Theiler,91]. Recently, new stationary stochastic signals with power–like power spectra yielding apparent finite correlation dimension, as well as some tests towards discerning them from chaotic signals have been discussed [Provenzale et al,92]. It is important to note that such 'finite–dimensional' stationary noises become (slowly) space–filling when the length of the series tends to infinity. In this report we also show a family of stochastic processes exhibiting this finite–dimensional behaviour (see chapter **3**).

A true colored (power–like) spectrum cannot be rendered by chaotic motion on a low-dimensional strange attractor. Unfortunately, a number of dynamical systems have been documented (see the references in [Theiler,91]) which produce finite–length orbits with colored spectra over a wide range of scales. However, a main point in [Theiler,91] is that a 'good' dimension algorithm (correctly applied) should not be fooled by signals with colored spectrum. Some useful recommendations addressing this issue are given in the paper by Theiler. In chapter **4** we propose a test, which stems from the theory of Liapunov exponents, that may be able to unmask finite–length noises of false finite–dimensional behaviour.

2.2.2 Practical Determination of the Correlation Dimension.

The first problem we are concerned with is to detect the presence of chaotic dynamics in given time series. For that aim, we look for a significant qualitative behaviour of

the correlation integral estimates in increasing embedding dimensions. If a chaotic mechanism is identified by means of rough but robust estimates, a second task is to obtain accurate estimates of the value at which correlation dimension saturates. This second issue has deserved most attention from the area of dynamical systems (where the first problem obviates), whereas the former has usually been pursued by field experimentalists.

We will concentrate on some general aspects of the qualitative analysis pertinent for chaos searching in our series, and will not worry too much about error estimates and other quantitative issues. In appendix **A**, however, we will report on some interesting results on the estimate of the correlation integral by least–squares.

A number of considerations about dimension estimation both in time series analysis and (also) in a general context are given in [Grassberger et al,91] and [Theiler,90] respectively, as well as *ad hoc* references for specific problems. [Grassberger,88], [Holzfuss & Mayer-Kress,86], [Mayer-Kress,87], [Caswell & Yorke,86], [Theiler,90b], [Ramsey & Yuan,89] are just some of the number of papers focusing on the important error-estimation problem in dimension calculation.

Extraction of the slope from a log-log plot.

Let m be fixed, and denote the m–dimensional correlation dimension $\nu_m(\mu)$ by ν. Assume that the limit in (2.16) exists so that the scaling law $C(r) \sim r^\nu$ holds as r tends to zero. Since the limit in (2.16) is not numerically attainable, the naive attempt to compute ν (in some embedding dimension) is to find

$$\nu \approx \frac{\log C(r)}{\log r} \tag{2.18}$$

for the smallest available r (i.e. the smallest r for which the correlation integral has been computed). Assume typically that the relation

$$C(r) = C_0(r) r^\nu \tag{2.19}$$

holds for small r's, where $C_0(r)$ depends on r and it is such that

$$\lim_{r \to 0^+} \frac{\log C_0(r)}{\log r} = 0. \tag{2.20}$$

Now, since the second term in the following estimate $\hat{\nu}$ of ν

$$\hat{\nu} \approx \frac{\log\left(C_0(r)r^{\nu}\right)}{\log r} = \nu + \frac{\log C_0(r)}{\log r}, \tag{2.21}$$

vanishes logarithmically slowly, equation (2.21) may not be a reliable estimate for ν. This situation persists even in the 'easy' case of self-similar fractals [Barnsley,88], where the function C_0 is usually constant. Observe that this estimate problem typically appears when any other quantity has to be obtained from an asymptotic logarithmic scaling law (e.g. box-counting dimension, pointwise dimension, etc.), so that what is said below also applies.

The usual method is to extract ν from a log-log plot of the correlation integral $C(r)$ vs. r over the range of scales of r for which $C(r)$ was computed. The slope of this 'curve' gives an estimate of ν for small values of r. In practice, one looks for a range of scales, as wide as possible, where a scaling holds. There is ample literature about the way this is best achieved, particularly concerned about the meaningful range (r_0, r_1) where the statistics provides a scaling law (i.e. ideally a straight line in the log-log plot). It is clear that such a range will be bounded above and below due to the finiteness of the data series and the boundedness of the attractor respectively. The maximum range of validity of the scaling law goes from $C(r, N) = 2N^{-2}$ to $C(r, N) = 1$, that is an order of $2\log N$ over the range of $\log C(r, N)$ in the log-log plot.

We do not discuss here the issue of the minimum length of the series for obtaining a meaningful dimension estimate, since the size of the data series we consider in this report seems to be large enough to proceed with the analysis. For papers discussing this subject see [Theiler,90] (and references therein), [Ruelle,90] and [Essex & Nerenberg,91]. Observe that the finiteness of the sampling size implies that the scaling range is bounded. It is also important to note that many of the systematic effects encountered in the correlation integral analysis would disappear in the limit $N \to +\infty$.

The criterion of fitting a chord through sucessive points of the plot (see for instance [Theiler,90]) seemed to produce robust estimates of ν when we checked it on some self-similar sets and known attractors, both for box-counting and for correlation dimension. The procedure is as follows. Let $\mathbf{r} = (r_1, r_2, \ldots, r_k)$ be a

fixed vector of radii $r_1 > r_2 > \ldots > r_k > 0$ for which the finite correlation integral $C(\cdot, N)$ has been computed. Our code typically works with 600 radii thus supplying 600 values of the correlation integral. An estimate of ν is then obtained by taking the local slope

$$\hat{\nu}_i(N) \approx \frac{\log C(r_{i+1}, N) - \log C(r_i, N)}{\log r_{i+1} - \log r_i}, \tag{2.22}$$

for $i = 1, 2, \ldots k-1$, and then looking for a *plateau* in a plot of local slope vs. radius. In order to enhance the robustness of the estimate we recursively average the values $\hat{\nu}_i(N)$ so that a new local slope estimate for ν is obtained at each r_i by averaging with the previously obtained values.

In order to understand the outline of the behaviour of the family $\{C^m(r)\}_m$ when m increases, all the log-log plots of successive m–dimensional correlation integrals are plotted to get the qualitative picture. What is relevant from the point of view of a qualitative analysis is whether the family $\{\nu_m\}_m$ of slopes estimated from the 'lines' $\log C^m(r) = \hat{\nu}_m \log r$ stabilizes to some finite value $\hat{\nu}$ or, on the contrary, sketches an increasingly non–convergent sequence of values. This behaviour is better sketched in a plot of ν_m vs. m, once all the values ν_m has been estimated from a significant scaling region (that can be different in each embedding dimension). Particular estimates will be extracted by the fitting chord estimate (2.22) when necessary.

Dynamical Correlation and Lower Cut-off Time.

The scaling law (2.16) for $C(r)$ reflects the fact that for a set of dimension ν, the number of points expected to be found in a random ball of radius r through the set behaves like r^ν. This should be taken into account when the set is obtained by a Ruelle–Takens reconstruction of a time series. This is because a long autocorrelation time combined with a short sampling time causes pairs of points to be considered as being close just because they are close in time. It has been documented [Theiler,86] that this unfair correlation sum produces an anomalous 'shoulder' in the correlation log-log plot. To avoid this problem it is recommended, rather than wasting data by increasing the sampling time [Grassberger et al,91], to consider for the sum $C(r, N)$ just those points within a distance r in phase space but more separate in time than a *lower cut-off time* τ_0. Typically τ_0 is chosen as the *correlation time* of the series.

The new correlation integral will be

$$C(r,N,T_0)=\frac{2}{(N+1-T_0)(N-T_0)}\sum_{j=T_0}^{N-1}\sum_{i=0}^{N-1-j}\Theta(r-\|x_i^{(m)}-x_{i+j}^{(m)}\|), \qquad (2.23)$$

where T_0 is the least integer greater or equal than $\tau_0/\Delta t$ and Δt is the sampling time. The code we have developed admits the cut–off delay T_0 as an input.

2.2.3 Reading Features of the Observed System from the Correlation Integral Curves.

There is a variety of interesting aspects of the signal nature that can be guessed from a correlation integral log–log plot with structure. Many of these remarks are collected in [Theiler,90]. Since a clear–cut hallmark of chaos is not expected to be found in field (noise–corrupted) signals, this deeper analysis of the correlation integral behaviour, beyond the search for a scaling region, is highly recommended. Notice that the 'shoulder' effect commented in the last paragraph is an example of structured correlation plots. We point out below one of these 'pathologies' in the correlation integral scaling, since it might concern our analysis. We also comment on a feature in the correlation plot that we have found in our analysis, and which may appear in many other contexts.

Systems coupling. The "knee" effect.

A signal measured from 'naturally' intricate phenomena (such as climate, developed turbulence, future markets, etc.) is probably best understood as the output of different interacting processes, experimental noise being the most obvious. When the processes are distinct enough, the coupling can be inferred from the $\log C(r)$ vs. $\log r$ plot. If two different ranges (say, (r_0,r_1) and (r_1,r_2)) are found so that a slope $\hat{\nu}$ is estimated in the range (r_0,r_1) and a smaller slope $\hat{\nu'}$ is in the range (r_1,r_2), then two distinct processes might be working at amplitudes below and above r_1.

The simplest case is that of two non-interacting (say, deterministic) systems, 1 and 2. Consider an observable $u=u_1+u_2$, where u_1 and u_2 are independent signals

from each component system. Assume further that the system of the signal u_1 is high–dimensional, while the system of u_2 is low–dimensional. If the amplitude r_1 of the signal u_1 is much smaller than that of u_2, then for values of $r > r_1$ we can only obtain statistical information about system 2. This will result in a small slope, $\hat{\nu}_2$ in the correlation plot. Instead, the amplitudes of the whole system are taken into account for $r < r_1$, which gives a larger slope $\hat{\nu}_1$, $\hat{\nu}_1 > \hat{\nu}_2$, as the correlation dimension of the system.

In many experimental settings a "knee" has been observed in the correlation plot, in such a way that the larger estimated slope $\hat{\nu}_1(m)$ coincides with the embedding dimension m whereas the smaller slope $\hat{\nu}_2(m) \simeq \hat{\nu}_2$ remains constant for sucessive values of m. This is a suggestion for low-dimensional dynamics corrupted by a space-filling process, what is consistent with the hypothesis of determinism (chaotic or not) corrupted by white noise of small amplitude. This effect is illustrated in section **C.7** for a Henon attractor with an added white noise of amplitude 0.02.

A, say one-dimensional, random walk can also produce a "knee" structure in the correlation plot since the random walk prevails at small scales (so that $\hat{\nu}_1(m) = m$ upon this range) whereas a one-dimensional linear structure is appreciated at macroscopic scale (what implies $\hat{\nu}_2(m) \simeq 1$ at large scales).

Limited–precision data. The 'stair–shaped' correlation plot.

We have found in our analysis of SWH–series correlation plots with a sort of stepped structure such as that exhibited in figure C.26. This effect does not allow us to proceed with the analysis, since no scaling region is available to compute the correlation dimension. The type of stair structure seen in fig C.26 may be due to data sets measured with an apparatus of limited precision. This limitation, when observing stationary phenomena with a moderate range of variation with respect to the order of precision of the apparatus, means that a large amount of different states of the process are measured as being the same. The implications of this fact on the correlation integrals is clear: the same number of neighbouring pairs are found for a wide range of different radii, thus giving the same value of $C(r)$ upon this range. This causes the stair–shape in the correlation plot. In section **2.3.2** we address this

problem and propose a solution to analyze with confidence this type of data.

2.2.4 The Grassberger's Box–Assisted Algorithm for the Correlation Dimension Computation.

The algorithm implemented for the computation of the correlation integral is described in detail in this section. The base code of Grassberger for the correlation dimension calculation is that in [Grassberger,90], whose basic ideas are presented below.

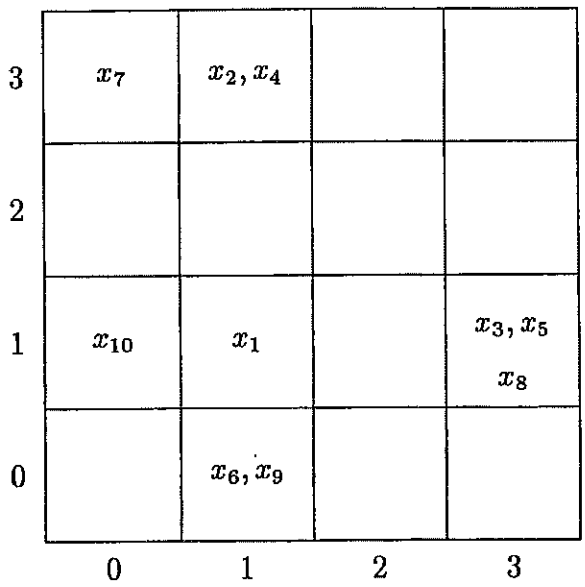
Let N be the length of the data series. The number of calculations to compute all the pair distances within a given radius $\varepsilon > 0$ is of the order of N^2. Since the correlation integral must be evaluated for a range of different radii and for several embedding dimensions, a naive implementation is very expensive in terms of CPU-time. The box-assisted technique devised by Grassberger allows us to analyze long series for a wide range of reconstruction dimensions in a reasonable time.

Box-assisted technique.

The core idea consists of storing in a matrix the geometrical disposal of the 2-dimensional histories of the series, $A_2 = \{x_n^{(2)}\}_{n=1,\ldots,N}$. Recall we consider the distance of the maximum in $\mathbb{R}^m$ (the correlation integral does not depend on the metrics [Brock,86]). Since we are only concerned with those pairs of points closer than ε, we first construct a grid of L^2 disjoint (e.g. dyadic) squares of side ε covering the data set A_2. An $L \times L$ matrix and a (N-2)-dimensional vector, called by Grassberger BOX and LLIST respectively, are constructed as follows. All the entries both of the matrix and the vector are set to zero and the placement of each data point in the grid is sequentially stored by setting $\mathrm{BOX}(i,j) = k$ if the k-th point $x_k^{(2)}$ of the embedded series is the first one in the cell (i,j). Assume that $m_k - 1$ points have been scanned, and that the point $x_{m_k}^{(2)}$ lies on the cell (i,j). If $\mathrm{BOX}(i,j) = m_0$ then the cell (i,j) is already occupied and the point $x_{m_0}^{(2)}$ was the first to lie on it. To get the complete history of the points that lie on cell (i,j) the

m_0-th element of the vector LLIST is considered. If LLIST(m_0) = 0 then we set LLIST(m_0) = m_k, which means that $x^{(2)}_{m_k}$ was the second point to settle in cell (i,j). If, on the contrary, LLIST(m_0) = m_1 we tune the first j such that LLIST(m_j) = 0 and set LLIST(m_j) = m_k. Therefore the matrix BOX stores the information of the empty boxes of the grid (those such that BOX$(i,j) = 0$) and of the ones occupied (BOX$(i,j) \neq 0$), whose sequencial occupation history can be recovered from the vector LLIST (i.e. {LLIST(BOX(i,j)),LLIST(LLIST(BOX(i,j))),...} is the series of index of the points occupying the cell (i,j), ordered according to the arrival time). Now, to find all the pairs within a distance ε, we have to check for each point just the distances to those points on the same cell and to those on the eight adjacent boxes.

To illustrate the construction of the matrix BOX and the vector LLIST, consider a ten-point 2-dimensional data set whose geometrical arrangement in a grid of 16 cells in the plane is that in figure **2.1**.



$$\text{BOX} = \begin{pmatrix} 0 & 10 & 0 & 7 \\ 6 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \end{pmatrix}$$

LLIST= (0,4,5,0,8,9,0,0,0)

Figure 2.1.

The index history of occupation of, say, cell (3,1), is recovered from BOX(3,1) = 3, LLIST(3) = 5, LLIST(5) = 8. Observe that LLIST(8) = 0 means that x_8 was the

last point to lie in box $(3,1)$. In general the index history of the box (i,j) is given by

$$\{LLIST^n(BOX(i,j))\}_{n\in\mathbb{N}},$$

where $\mathrm{LLIST}^n(\cdot)$ denotes $\mathrm{LLIST}(\mathrm{LLIST}^{n-1}(\cdot))$, $n \geq 1$, and LLIST^0 is the identity.

The algorithm.

Let $R = \{r_i,\ i = 0,1,\ldots,l\}$ be the collection of radii in decreasing order for which the correlation integral is being computed. Let $m_0 \geq 2$ and $M = \{m_i :\ i = 0,\ldots,q\}$ be the set of scanned embedding dimensions. Let BOX and LLIST be the matrix and vector constructed by the box-assisted technique with $\varepsilon > r_0$. Now the cell of a given $x_{i(m)}$ is found (see below), so that the points x such that $d(x, x_i^{(m)}) < r$ for some $r \in R$ must lie in the same box as $x_i^{(m)}$ or in some of the adjacent boxes. For $x_i^{(m)}, x_j^{(m)} \in A_m$, let k be the first index such that $d(x_i^{(m)}, x_j^{(m)}) \equiv d < r_k$, Then clearly $d > r_p$ for $p > k$. This observation implies that a single evaluation of each distance in $\{d(x_i^{(m)}, x_j^{(m)})\}_{i\neq j}$ gives at the same time all the pairs $(x_i^{(m)}, x_j^{(m)})$ within every radius r in the set R. This information is stored in a matrix MM of dimension $q \times l$ defined by

$$\mathrm{MM}(m,i) = \sharp\{(i,j) :\ d(x_i^{(m)}, x_j^{(m)}) < r_i\},\ \ (m, r_i) \in M \times R.$$

Observe that the computation of the distances for all the dimensions m in M simultaneously allows us to reduce drastically the number of comparisons, since two points far apart more than ε in dimension m will be at least that far in dimensions greater than m. Thus those pairs farther than ε in dimension m do not have to be compared again in upper dimensions.

The calculations can be further reduced by constructing the matrix BOX and the vector LLIST at the same time as the distances are being computed. So the algorithm begins with the placement of the point $x_1^{(2)}$ in its corresponding cell of the grid and its recording in the matrix BOX. The distance $d(x_1^{(2)}, x_2^{(2)})$ is computed only if $x_2^{(2)}$ lies on the cell of $x_{(1)}$ or on some of the adjoining cells. In general, once the cell (i,j) of the point $x_n^{(2)}$ has been recorded in the matrix BOX or the vector

LLIST only those points $x_p^{(m)}$ in the cells (i,j), $(i,j\pm 1)$, $(i\pm 1,j)$ and $(i\pm 1,j\pm 1)$ have to be checked . The matrix BOX and the vector LLIST will not be finished until all the data points have been scanned. Note that in this way each pair (i,j) is considered just one time (i.e. (j,i) will not if $i<j$) so that the calculations are a half of those in the naive algorithm.

Grassberger considers the scale range given by $R=\{r_i=2^{-i/2},\ i=0,\ldots,127\}$. Then the first index k such that $d\equiv d(x_i^{(m)},x_j^{(m)})<r_k$ is given by $k=\mathrm{int}(-2\log_2 d)$. We consider in our code all the radii of the form $2^{-i/\mathrm{ZE}}$ where $i=1,2,\ldots,600$, and ZE is an input of the program that depends on the user's necessities of scanning smaller resolution scales. This actually is related to the available amount of data.

If the bounded region of the phase space where the set A_2 lies is not known, or if the grid covering the set requires a large number of squares of side ε, the dimension of the matrix BOX will be too large and severe problems with computer memory can arise. Grassberger solves this problem by fixing the dimension L of the matrix box so that the grid occupies the square $Q=[0,\varepsilon L]\times[0,\varepsilon L]$. The whole plane is then rolled over the torus Q. Now, the cell (i,j) where a point $x_k^{(2)}=(a,b)$ lies is given by $i=\mathrm{int}(a/\varepsilon)\bmod L$ and $j=\mathrm{int}(b/\varepsilon)\bmod L$. The adjacent cells are (i_1,j_1) with $i_1=i-1,i,i+1$ and $j_1=j-1,j,j+1$ provided that $0\le i_1,j_1\le L-1$, and taking the remainder of the division by L otherwise.

Though specially addressed to the efficient calculation of the correlation integrals, the compact code of Grassberger can be easily modified if some additional information is needed. Moreover, some of the smart ideas in Grassberger's code can still be efficiently applied to the computation of the pointwise dimension. Specifically, a histogram of the number of neighbors closer than a distance r to each point of the data set embedded in dimension $m\in M$ can be obtained by minor changes in the source code. To that aim, the matrix BOX and the vector LLIST must be completely generated. Then a matrix denoted by $\mathrm{IP}(m,k)$ stores the number of points closer, in dimension m, than a distance r_k to each point. This information allows us to construct the matrix $\mathrm{IFREC}(m,k,i)$ which records the number of points having i points within a distance $d<r_i$. Note that the pointwise dimension distribution can also be estimated using the box–assisted technique.

2.3 Application to Wave Height Time Series.

In this section we thoroughly analyze the correlation integral behaviour of several time series monitoring significant wave height from the Spanish sea swell. In particular, the work is devoted to significant wave height time series of the *mar Cantábrico* (Cantabrian sea), obtained from measurements registered by oceanographyc buoys moored in the north coast of Spain. Significant wave height is one of the most widely used parameters to describe the evolution of the sea wave height. This section is structured as follows. In the subsection **2.3.1** the methods used for obtaining the time series displaying significant height is briefly explained. In **2.3.2** the correlation integral analysis of the different time series under consideration is addressed. What seems to be a new method to perform a trustworthy correlation dimension analysis on truncated–observation data is also presented in this subsection. Conclusions from the developed testing will be reported in chapter five.

2.3.1 Obtaining Significant Wave Height Series from Raw Data of Sea Elevation.

The input for constructing the significant height time series are raw data measurements of sea elevation. Recall that we abbreviate 'significant wave height' by SWH. A scalar buoy moored on the coast of Gijón (a northern town by the Cantabrian sea) records the sea elevation each 1/2 second for 45 minutes approximately to obtain 5120 scalar data and then stops the recording. Three hours after having begun the previous registration the recording process starts again. A single significant height value is obtained from the 5120 data so that the final output for the SWH–series is lagged 3 hours.

There are two different methods for obtaining SWH–time series from raw data measuring sea elevation, namely the statistical and the spectral method.

Statistical Method.

Sea elevation data are positive or negative (with zero mean). Individual waves are built up from the sea elevation data as follows. The beginning of the first wave is taken from the first change of sign from negative to positive values in the crude elevation series, whereas the end of the wave is given by the following ascending change of sign. A sequence of individual waves is obtained in this way from the 5120 data. The height of each wave is computed by subtracting the absolute value of the lowest (negative) elevation to the highest (positive) elevation recorded during the wave–period. All these raw wave height values are (decreasingly) ordered according to their heights, and then the average is computed from the highest third of the values. This averaged parameter is taken to be the SWH–value in the 3–hours period, which we denote by H_S.

Spectral Method.

The raw elevation data form a real–valued time sequence of N points ($N = 5120$ in our case) with a sampling time of ΔT seconds ($\Delta T = 0.5$ for us). A Fast Fourier Transform (FFT for now onwards) is performed so that the spectral density $S(\omega)$ is computed from the equation

$$S(\omega_k) = \frac{1}{2\,\Delta\omega} \mid C(\omega_k) \mid^2,$$

where $\{C(\omega_k)\}_{k=0,\ldots,N/2}$ is the complex-valued series of the Fourier coefficients (output of the FFT–routine), $\omega_k = k\,\Delta\omega$ are the sampling frequencies, and $\Delta\omega = (N\,\Delta T)^{-1}$ is the sampling frequency interval. Now, the 0th–order moment of the spectrum is computed

$$M0 = \left(\frac{N}{2} + 1\right)^{-1} \sum_{k=0}^{N/2} S(\omega_k),$$

and the SWH–value for a 3–hours period is taken to be

$$H_{M0} = 4(M0)^{1/2}.$$

It is known that both SWH–parameters, H_S and H_{M0}, give quite similar values, H_{M0}–values being 5% above H_S–values. For a detailed discussion on the methods outlined above see [Conde & Calderón,86].

We have developed the analysis for several H_S–series, as well as H_{M0}–series, of different recording–dates and lengths. We will concentrate on two series, one obtained from the statistical method, and the other one obtained from the spectral method. The H_S–series Gi_S has 24,098 height data obtained from the statistical method in the period between January 1985 and December 1994. The SWH–series Gi_{M0} is a H_{M0}–data series with 23,102 points obtained from the spectral method that runs through the same period as Gi_S. The Gi_{MO} series and the Gi_S logarithmic series can be seen in figure C.24. The sampling time between sucessive measurements is 3 hours in both series. There are missing data in both series, and that is why they have different length. This lack of data is not important in order to perform the correlation analysis so that the gaps in the series were not filled by some autoregresive method. Furthermore, due to the noisy structure introduced by the measure apparatus in the recording of the elevation series (see **2.3.2**), the missing data seem not to be relevant for the analysis. The same is not true for the Liapunov spectrum computation, so that the series has to be manipulated in order to obtain a segment–series without gaps (see chapter **4**).

2.3.2 Correlation Dimension Analysis of SWH–Series. A Method for Analyzing Data Observed with Limited Accuracy.

Let $\nu(m)$ denote the estimated slope from the plot of the correlation integral ($\log C(r)$ vs. $\log r$). We compute in section **5** of the appendix **C** the correlation dimension of our testing attractors as well as of space–filling stochastic processes. The correlation plots as well as the plots of $\nu(m)$ vs. m, up to a value of $m = 30$, computed from reconstructed trajectories of x–coordinate series with 30,000 (15,000 for the Henon system) data points observed in the attractors of Henon, Lorenz, and Rössler systems can be seen in figures C.14, C.15 and C.16 respectively. In contrast, the same plots for a series with the same data points from a gaussian process (generated

from a routine of the I.M.S.L.) exhibit a very different structure, giving increasing estimations of $\nu(m)$ with respect to the dimension of the space where the data set is reconstructed (see figure C.17). The space–filling nature of the white noise is clearly exhibited in the $\nu(m)$ vs. m plot, where $\nu(m) \simeq m$ up to dimension $m = 12$ or 13. From these dimensions up some estimation problems arise and a scale range to compute $\nu(m)$ is no longer available. We use these examples as a testing bench for the developed code that also illustrates the type of behaviour that can be expected when the correlation dimension is computed. We proceed to analyze the series Gi_S and Gi_{M0}. The delay reconstruction of the series for embedding dimensions 2 and 3 and for different delay times is shown in figure C.25. Notice that the input for the correlation computation program is the scalar time series, so that no reconstruction need to be done.

The crude correlation integral for the series Gi_S is shown in figure C.26. The log–log plot exhibits an anomalous 'stair–structure' that does not show a range where a clean slope can be estimated. In fact, no scaling law can be asserted to exist from such a plot. We know that this type of 'stair–shaped' correlation plot is caused by a large amount of equal data values. In fact, the number of different H_S-data in the Gi_S series is just 586 (from a number of 24,098 recordings). This obviously implies that the correlation sum $C(r)$ will be equal for a large number of different (but close) values of radii r, which introduces the structure in the correlation plot that is appreciated in C.26 (see also figure C.29).

Notice that the stepped–structure above is the effect expected to be found in the analysis of experimental data obtained by finite–precision measuring apparatus. The limited resolution of the observing mechanism introduces an experimental noise that rounds off the measured data so that different (though similar) realizations of the underlying process cannot be distinguished in practice. In our case, the buoy has resolution up to the order of centimeters, which implies that distinct 'sea–state evolutions' are spuriously observed as the same by the buoy measuring mechanism. Notice that this problem, which can be called of *limited–precision data*, may appear in some other stationary phenomena (e.g. economic or financial series) where exact data are available but the data values are very similar. This kind of data can be regarded as the true real data with an added noise of truncation (or round–

off), which may be modelled by a stochastic process consisting of independent and identically distributed (i.i.d.) random variables with a uniform distribution whose amplitude is the precision of the apparatus of observation.

We introduce the following method for analyzing *data of round–off observation.* Let ρ denote the precision of the observer. First, several synthetic series are created from the truncated series by adding, up to resolution scale ρ, different realizations of an i.i.d. process with uniform probability distribution in the interval $(-\rho/2, \rho/2)$. Secondly, the correlation analysis of the new series is developed. The artifact of adding up the small-amplitude noise destroys the stair–shape in the correlation plot (this can be seen in figure C.27 for the Gi_S–series). Lastly, correlation integrals and dimensions can be estimated (within an statistical significance) from the series ensemble where the analysis was performed. Of course, the results are not significant within the radius range $(0, \rho)$ that corresponds with the amplitude of the added noise. This range is visually displayed in the correlation plot (when the process is not space–filling) by a 'knee' structure separating the noise log-scale range $(0, \rho)$ from the log-scale range where the slopes can be significantly estimated. This effect was commented on in **2.2.3**, and can be appreciated in figure C.23 for the Henon attractor. Notice that the dimension can be correctly estimated out of the noisy range.

The proposed method to deal with the height round–off data can be understood as follows. We recovered an ensemble of different sea state evolutions that may have generated the height signal recorded by the buoy. This can be seen as the process inverse, in a sense, with respect to the measuring mechanism. Obviously, all the elements of the ensemble share the properties of the height signal, that will be taken out by the statistical analysis of the ensemble.

Notice that the method above also applies to *data of truncated observation*, which is other form of limited–precision observation. In this case, if the precision of the observer is given by ρ, the false series are made by adding white noise of amplitude $(0, \rho)$. The amplitude of the noise is also of the order of the precision of the apparatus, and the results are thus reliable in the corresponding scaling range.

The results of the technique above when applied to the series Gi_S can be seen in

figure C.28. Since the precision level of the buoys is of the order of 0.01 (centimeters), we have taken $\rho = 0.005$ in our case, and have created ten different spurious series. The results were the same for all the plots within the valid range. A correlation dimension plot of the ensemble of ten synthetic series can be seen in C.28. Notice that a true slope can be significantly estimated in the new plot up to $m \simeq 15$. Observe also that a knee cannot be distinguished in this plot. This is explained by the fact that the series Gi_S behaves as noise at all scales. The estimated dimension $\nu(m)$ does not reach the value of the embedding dimension m, but this may be explained by a lack of data. It is known that the number of points needed to fill the embedding space increases exponentially with m. What is significant in this case is the non–stabilized behaviour of the sequence $\{\nu(m)\}_m$ as m increases, which does not allow us to accept the SAH for the studied SWH–series.

Taking logarithms is a standard artifact to make a series stationary [Box & Jenkins,76]. Since this will be occasionally done to our series in order to perform some statistical tests elsewhere, we have also worked with series that are logarithms of H_S, which we will called $\log Gi_S$. We have computed the correlation dimension of two logarithmic series created from the Henon system, the result of which is reported in figure C.21. Note that the results are the same obtained for the series without logarithms, as was expected to happen due to the smoothness of the logarithmic function (see **A.2.2**). However, the compression effect of the logarithm on the data set might produce qualitative differences in the plot (for large values of radii). We have plotted in figure C.20 the 2–histories of two logarithmically transformed x–coordinate series from the Henon dynamics. Notice that the scale is different in the top figure of C.20 from that in the true dynamics C.1. The bottom plot in figure C.28 shows that the logarithmic H_S–series, $\log Gi_S$, gives a correlation analysis truly similar to that of the series Gi_S. Notice that this also provides robustness to the estimated behaviour of the H_S–series.

We developed the same work program for the series Gi_{M0}, and no difference was found with respect to the analysis of the series Gi_S. For the sake of completeness, we have included some of the results for this series in C.29.

Chapter 3

Testing Entropy. The K_2–Entropy

We address in this chapter another useful invariant under measure-isomorphism for chaos testing. More precisely, the *Kolmogorov–Sinai (*KS*) entropy* (or measure-theoretic entropy) is a measure of the *randomness* present in a system activity, read as the information gained when an experiment is performed and its output is given. In section **3.1.1** we present in detail the concept of KS entropy, together with a discussion of the role to be played by entropy as a tool in chaos detection. In a second section, the concept of K_2–entropy, more useful for practical purposes, is introduced with a detailed explanation of its efficient implementation from the code developed for the computation of correlation dimension (see chapter **2**). The last section concerns the study of chaotic behaviour in the SWH–series, by means of using spectral techniques along with the testing program developed so far. As a by–product of our research, we show a family of pure random noises that apparently fool the tests devised in chapters **2** and **3**. The main facts about entropy theory are nicely presented in [Walters,82] and [Billingsley,78] (the latter in connection with information theory).

3.1 Measure–Theoretic Entropies.

3.1.1 Kolmogorov–Sinai Entropy.

It has been argued in chapter **1** that sensitivity to initial conditions (SIC) in chaotic systems generates uncertainty in the future state of the system as soon as the information about the initial state vanishes. A system enjoying SIC creates information, in the sense that it removes uncertainty when it informs about its future state. Both the uncertainty (*before* the experiment is done) and the information (*after* the experiment has been performed) measure the degree of stochasticity in the experiment.

The situation is quite the same as that of a k-sided die–tossing with probabilities $\mathbf{p} = \{p_1, \ldots, p_k\}$, whose entropy is given by $H(\mathbf{p}) = -\sum_{i=1}^{k} p_i \log p_i$ (with the convention that $0 \log 0 = 0$). The value $H(\mathbf{p})$ gives a plausible measure of the amount of uncertainty of the die $\mathbf{p}$: $p_i = 1$ for some i implies $H(\mathbf{p}) = 0$, and there is no uncertainty about the outcome; whereas $p_i = k^{-1}$ for all i gives the uncertainty about the output of the 'most random' die, $H(\mathbf{p}) = \log k$ (which can be shown to be a maximum over the different probability configurations for the sides of the die). We give below the construction of the entropy invariant of a μ–measure–preserving transformation T on the compact support of μ.

The entropy of the measurable mapping f preserved by the measure μ is defined in three steps. First, the *entropy of a measurable finite partition* $\xi = \{A_1, \ldots, A_k\}$ of the support of μ is defined by

$$H(\xi) = -\sum_{i=1}^{k} \mu(A_i) \log \mu(A_i), \tag{3.1}$$

so that ξ splits all the possible outcomes in the phase space from an observation of the system state into k 'faces' and $H(\xi)$ measures the uncertainty removed (or information obtained) when the experiment represented by observing the system state is performed. Observe that this is the situation described above for the k-sided die. In fact, definition (3.1) also applies in this context [Renyi,70]. Therefore, from an information entropy viewpoint, 'tossing a die' must be considered as a more ramdom experiment than 'tossing a coin' when the partition whose elements are the

atoms of the probability measures are considered.

Secondly, the *entropy of f relative to ξ* is defined by considering the combined experiment of performing the original experiment indefinitely along the time evolution. For that aim, the partition $\xi^{(n)}$ generated by ξ after n time steps is defined as the collection of the sets

$$A_{i_1} \cap f^{-1}(A_{i_2}) \cap \ldots \cap f^{-n+1}(A_{i_n}), \;\; i_j \in \{1, 2, \ldots, k\} \tag{3.2}$$

where $f^{-j}(A_i) = \{x \in \text{supp}\mu : \; f^j(x) \in A_i\}$. The entropy of f with respect to ξ is then defined by

$$h(f, \xi) = \lim_{n\to\infty} (H(\xi^{(n+1)}) - H(\xi^{(n)})) = \lim_{n\to\infty} \frac{1}{n} H(\xi^{(n)}), \tag{3.3}$$

that represents the asymptotic averaged information obtained by unit of time when the original experiment is infinitely repeated. This can be considered as the *statistical* version of the entropy (3.1) of a single trial of the experiment. It also makes sense for stochastic processes, where f is just the time-passing mapping or *shift* transformation. We will turn out attention back to this point later.

Lastly, the *entropy of the pair* (μ, f) is defined as

$$h(\mu) \equiv h(\mu, f) = \sup_{\xi} h(f, \xi), \tag{3.4}$$

where the supremum is taken over all the partitions of $\text{supp}\mu$. $h(\mu, f)$ is therefore the maximum averaged information obtained per unit time when the same experiment is infinitely repeated.

The supremum in (3.4) can be avoided if ξ is a *generating* partition (this basically means that the sets in (3.2), $n \in \mathbb{N}$, generate the σ-algebra of the measure space). This is the content of the famous *Kolmogorov–Sinai* theorem. On the other hand, it is worth noting that $h(f, \xi) > 0$ for some chosen partition ξ is a sufficient condition for randomness in the process $(\text{supp}\mu, f)$. However, note that the supremum in (3.4) is crucial for concluding that two discrete Bernouilli processes with two (coin) and six (die) equally probable states have different entropies, respectively given by $\log 2$ and $\log 6$. Kolmogorov proved this fact along with the isomorphism-invariance of the entropy to show that two different Bernouilli processes with discrete state space are not isomorphic (the invariant h was actually introduced to solve this problem).

The 'purest' of the classical stochastic processes, namely throwing a fair die, has positive finite entropy $\log k$. Moreover, given an integer k, it is easy to construct a one–dimensional discrete deterministic system (F, L^1) (with the Lebesgue measure L^1) such that $([0,1), F, L^1)$ is isomorphic to the discrete uniform Bernouilli process with k states and thus with entropy $\log k$ (see e.g. chapter **1**). This fact connects with the profound problem that really lays the foundations of this research area: to what extent a stochastic process is representable as a dynamical system, and what advantages can be obtained from such representability.

It is known that $h(\mu) = 0$ for an ordered (non-chaotic) deterministic system while $h(\mu) > 0$ is an indication of randomness [Walters,82]. It has been suggested that $0 < h(\mu) < +\infty$ is a *sufficient* condition for chaos since it is said that random systems give $h(\mu) = +\infty$ [Grassberger & Proccacia,83b], [Provenzale et al,91]. This last assertion must be expected to hold only for some classes of space–filling stochastic processes. In fact, any process of independent random variables with absolutely continuous distribution can be shown to yield infinite entropy [Renyi,70]. Certain 'colored' noises, however, yield positive and finite K_2–entropy estimates [Provenzale et al,91] (see below for the definition of the K_2–entropy). We also give in **3.3.2** some examples of pure (without structure in their spectra) stochastic processes giving finite estimates of K_2–entropy.

From what is said above the capability of $h(f)$ for discerning between chaos and stochasticity in a practical setting seems to be limited. For a signal recorded on a dynamical system a positive and finite entropy estimate is a sufficient condition for the presence of chaos but the picture darkens for field-monitored signals. Therefore, a plausible experimental test for entropy will work best by rejecting the hypothesis that the system is ordered (zero-entropy) or space-filling (infinite entropy). Accepting the hypothesis of chaoticity requires more work, as for example that of characterizing those 'space-filling' noises. In any case, entropy is a valuable tool that should be added to a testing program on chaos. What follows strengthens this belief.

3.1.2 K_2-Entropy.

The Komogorov–Sinai entropy might be numerically evaluated [Eckmann & Ruelle,85, Grassberger & Proccacia,83b], but the task gets complicated in an experimental setting. Grassberger and Procaccia [Grassberger & Proccacia,83a] introduced the K_2-entropy that estimates the KS–entropy from below [Renyi,70], and is more easily computable than the KS–entropy. Much more than this is true, it can be estimated from the computer code developed for the correlation dimension. The K_2-entropy, one of the Renyi's q-entropies [Renyi,70], is defined by

$$K_2 \equiv K_2(\mu, f) = \sup_{\xi} \lim_{n\to\infty} \frac{1}{n} K_2(\xi^{(n)}), \tag{3.5}$$

where, for a partition ξ,

$$K_2(\xi) = -\log \sum_{i=1}^{k} \mu(A_i)^2 \tag{3.6}$$

and $\xi^{(n)}$ is the partition with those elements in (3.2). Again the supremum in (3.5) does not need to be computed for a generating partition. Notice that an estimate $K_2(\xi)$ for any partition ξ gives a lower bound for the true K_2-entropy.

A remarkable property of entropy is that it is naturally connected with the problem of *predictability* for the signal (2.9) when this signal was generated by some chaotic (or not) dynamical system. Since the KS-entropy $h(\mu)$ measures the averaged amount of uncertainty removed by a new observation of the state of the system, it is natural to consider $h(\mu)^{-1}$ as a measure of the available power of prediction from the system evolution. A sucessful prediction machinery for such a process should be built on the basis of the underlying nonlinear dynamics of the process.

The saw-tooth dynamics of chapter **1**, $f(x) = 2x - \text{int}(2x)$ with the Lebesgue measure, permits the role of the entropy to be illustrated in the prediction problem for a typical orbit $\{f^n(x_0)\}_{n\in\mathbb{N}}$, with $x_0 \in [0,1]$. Assume that $\tilde{x}_0$ is the noisy initial condition, that is $\mid \tilde{x}_0 - x_0 \mid = \delta \simeq 2^{-N}$. It is clear that after N time steps the initial (though tiny) noise is amplified up to the size of the phase space $[0,1]$, so that the (deterministic) predictive power is completely lost and the system becomes random. On the other hand, it is easy to see that $h(f) = K_2(f) = \log 2$ in this case. Therefore we have, as a clean formulation of the role of entropy in the prediction

problem,

$$N^*(\delta) \simeq -\frac{\log \delta}{\log 2} = c(\delta)h(f)^{-1}, \tag{3.7}$$

where $N^*(\delta)$ can be understood as the prediction time allowed by the available observation δ of the system state. Two remarks regarding equation (3.7) enlighten the SIC–phenomenon with respect to its connection with prediction theory. First, the capability of prediction improves if the capability of observation does. Second, which represents the dramatic feature of chaoticity, the only way to avoid the loss of the forecasting power in a finite time is the ability to observe the system state with infinite precision. Since this requirement cannot be fulfilled in experimental situations, chaos is suspect of being active, hidden in many natural phenomena.

A test that makes use of the correlation integral to evaluate the predictive power for the series is available [Grassberger et al,91]. There are also interesting methods based on local entropies and forecasters that try to identify regions/situations in phase space where prediction is feasible.

3.2 Implementation: Computation of the K_2–Entropy.

In practice, the K_2-entropy can be computed from the correlation integral so that the calculation is easy for the time series $\{u_i\}_{i=1,2,\ldots,N_p}$. The connection between the K_2-entropy and the correlation integral $C^d(\varepsilon)$ comes from exploiting the d–dependence in $C^d(\varepsilon)$ when d increases. We assume throughout this section that the SAH holds, and we use the notation from section **1.3**.

Let $d \in \mathbb{N}$ and $N \equiv N_d = N_p - (d-1)$. Let

$$A_d = \{x_n^{(d)} = (u_n, u_{n+1}, \ldots, u_{n+(d-1)})\}_{n=1,\ldots,N} \subset \mathbb{R}^d$$

be the d–dimensional Ruelle–Takens reconstruction of the series. We use the norm of the maximum in $\mathbb{R}^d$ and denote it by $\|\cdot\|_d$. Let M be larger than $2\dim\Lambda + 1$, so that $\hat{A}_M$ is diffeomorphic to the actual attractor Λ (see Takens theorem in chapter **1**). Let d be a positive integer and $\varepsilon > 0$. The (M+d+1)-dimensional correlation

integral can be written as

$$C^{M+d+1}(\varepsilon) = \lim_{N\to\infty} N^{-2} \; \sharp\{(n,m) : \| x_n^{(M)} - x_m^{(M)} \| < \varepsilon, \| \hat{f}(x_n^{(M)}) - \hat{f}(x_m^{(M)}) \| < \varepsilon, \ldots, \| \hat{f}^d(x_n^{(M)}) - \hat{f}^d(x_m^{(M)}) \| < \varepsilon\}, \tag{3.8}$$

since $\hat{f} \equiv \hat{f}_{M+d+1}$ behaves like a shift on the vectorized series (see **1.3** for the notation).

Now, the M-dimensional phase space is partitioned in a finite number of N_{box} disjoint boxes (e.g. dyadic M-boxes) of diameter $\varepsilon > 0$, $\mathcal{A} = \{B_i\}_{i=1,\ldots,N_{box}}$.

It is clear that

$$\sharp\{(n,m) : \| x_n^{(M)} - x_m^{(M)} \| < \varepsilon\} \geq \sum_i \sharp\{(n,m) : \; x_n^{(M)}, x_m^{(M)} \in B_i\}. \tag{3.9}$$

The sum in the M–correlation integral can be split as

$$\begin{aligned} &\sharp\{(n,m) : \| x_n^{(M)} - x_n^{(M)} \| < \varepsilon\} = \\ &= \textstyle\sum_i \sharp\{(n,m) : \| x_n^{(M)} - x_m^{(M)} \| < \varepsilon, \; x_n^{(M)}, x_m^{(M)} \in B_i\} + \\ &+ \textstyle\sum_{i<j} \sharp\{(n,m) : \| x_n^{(M)} - x_m^{(M)} \| < \varepsilon, \; x_n^{(M)} \in B_i, \; x_m^{(M)} \in B_j\} \leq \\ &\leq \textstyle\sum_i \sharp\{(n,m) : \; x_n^{(M)}, x_m^{(M)} \in B_i\} + k_d N_{box} (\sharp(N))^2, \end{aligned} \tag{3.10}$$

where $k_d = 3^d - 1$ and $\sharp(N) = \max_i \sharp\{n : \; x_n^{(M)} \in B_i\}$.

Now, if the asymptotics

$$\sharp(N) = o(N) \tag{3.11}$$

holds, we obtain from equations (3.8) to (3.11) that

$$C^{M+d+1}(\varepsilon) =$$

$$= \lim_N N^{-2} \sum_{i_1,i_2,\ldots,i_d} \sharp\{(n,m) : x_n^{(M)}, x_m^{(M)} \in B_{i_1},$$

$$\hat{f}(x_n^{(M)}), \hat{f}(x_m^{(M)}) \in B_{i_2}, \ldots, \hat{f}^d(x_n^{(M)}), \hat{f}^d(x_m^{(M)}) \in B_{i_d}\} =$$

$$= \lim_N N^{-2} \sum_{i_1,i_2,\ldots,i_d} \sharp\{(n,m) : (x_n^{(M)}, \hat{f}(x_n^{(M)}), \ldots, \hat{f}^d(x_n^{(M)})) \in B_{i_1} \times B_{i_2} \times \ldots \times B_{i_d},$$

$$(x_m^{(M)}, \hat{f}(x_m^{(M)}), \ldots, \hat{f}^d(x_m^{(M)})) \in B_{i_1} \times B_{i_2} \ldots \times B_{i_d}\}. \tag{3.12}$$

Assume now that the measure μ is ergodic. For any p large enough and any measurable $B \subset \mathbb{R}^p$

$$\begin{aligned} \mu(B)^2 = & \lim_N N^{-2} \left(\sharp\{n : x_n^{(p)} \in B\}\right)^2 = \\ = & \lim_N N^{-2} \sharp\{(n,m) : x_n^{(p)}, x_m^{(p)} \in B\}, \end{aligned} \tag{3.13}$$

holds μ–a.e., which applied to the boxes $B_{i_1} \times B_{i_2} \times \ldots \times B_{i_d} \subset \mathbb{R}^{Md}$ in (3.12) yields

$$C^{M+d+1}(\varepsilon) =$$

$$= \sum_{i_1,\ldots,i_d} \lim_{N\to\infty} N^{-2} \left(\sharp\{n : (x_n^{(M)}, \hat{f}(x_n^{(M)}) \ldots, \hat{f}^d(x_n^{(M)})) \in B_{i_1} \times B_{i_2} \times \ldots \times B_{i_d}\}\right)^2 =$$

$$= \sum_{i_1,\ldots,i_d} \mu\left(B_{i_1} \cap \hat{f}^{-1}(B_{i_2}) \cap \ldots \cap \hat{f}^{-d}(B_{i_d})\right)^2.$$

This means

$$C^{M+d+1}(\varepsilon) = \sum_{P \in \mathcal{A}^{(d)}} \mu(P)^2, \tag{3.14}$$

which is the key equality for the computation of the K_2-entropy from the correlation dimension.

Taking into account that the recording time in the original signal is Δt, we get, by substituing (3.14) in (3.6)

$$K_2(\varepsilon) \equiv K_2(\mu, \mathcal{A}) = -\frac{1}{d\,\Delta t} \lim_{d\to\infty} \log C^d(\varepsilon), \tag{3.15}$$

which gives the K_2-entropy associated with $\mathcal{A}$ from the correlation integral.

Errors of the order $1/d$ in (3.15) can be avoided by noting that

$$e^{K_2(\varepsilon)\Delta t} \sim \frac{C^d(\varepsilon)}{C^{d+1}(\varepsilon)} \tag{3.16}$$

holds as d tends to infinity. This gives that

$$K_2(\varepsilon) = \lim_{d\to\infty} K_2(d,\varepsilon), \tag{3.17}$$

where

$$K_2(d,\varepsilon) = \frac{1}{\Delta t} \log \frac{C^d(\varepsilon)}{C^{d+1}(\varepsilon)}. \tag{3.18}$$

It is known [Eckmann & Ruelle,85] that $\lim_{d\to\infty} \max\{|\ P\ |\colon\ P \in \mathcal{A}^{(d)}\} = 0$ is a sufficient condition for partition $\mathcal{A}$ to be generating. Therefore, the supremum in (3.5) can be dropped for such a partition. When this cannot be assured, the standard approach [Eckmann & Ruelle,85, Grassberger & Proccacia,83b] is to take

$$\tilde{K}_2 = \lim_{\varepsilon\to 0} K_2(\varepsilon) \tag{3.19}$$

as the K_2-entropy of the measure.

The numerical implementation for computing K_2 is now clear from equations (3.18) and (3.19). We estimate the K_2-entropy for several data sets in the following section.

3.3 Application to Significant Wave Height Time–Series.

We develop here the entropy analysis of the SWH–series introduced in chapter **2**. We also combine spectral analysis with correlation and entropy tests to detect a chaotic component in the evolution of the SWH–process recorded on the coast of Gijón. Recall that the series under study were defined in section **2.3.1**.

3.3.1 K_2–Entropy of SWH–series.

In this subsection, we report on the estimate of the K_2–entropy for the H_S–series according to the implementation explained in **3.2**. The computation of the K_2–entropy has not been put in practice in the literature so extensively as the correlation dimension has been. A systematic account of what can be found from this estimate procedure in the context of this report seems to be lacking. For these reasons we also include in this point some general considerations and computations for known examples that can illustrate the K_2–entropy analysis.

We recall that the estimate of the K_2–entropy is obtained from the the family of the correlation dimension estimates $\{C^m(\varepsilon)\}_{m,\varepsilon}$ by the formulae

$$K_2(m,\varepsilon) = \frac{1}{\Delta t}\log\frac{C^m(\varepsilon)}{C^{m+1}(\varepsilon)}. \tag{3.20}$$

and

$$K_2(\varepsilon) = \lim_{n\to\infty} K_2(m,\varepsilon). \tag{3.21}$$

In practice, some values of ε are chosen so that the correlation integral is still significant. To this end, we have prepared the code to give us, for every embedding dimension, the value of ε for which the correlation integral is zero. This may provide a criterium to choose adequate values of ε and an upper bound for the embedding dimensions so that the computations can be done with some confidence. We will give the output of the calculations of (3.20) by means of entropy–plots that represent the curves $K_2(\varepsilon)$ vs. ε for the chosen ε's.

We have first studied the behaviour of K_2–entropy estimates with respect to some of the data analyzed in this report. We have worked mainly with the Henon system, whose entropy–curves plot is given in figure C.31. A data series of the x–coordinate with 15,000 observation was used. The figure C.31 shows six entropy curves $K_2(\varepsilon)$ for the radii values $\varepsilon = 0.42,\ 0.17,\ 0.074,\ 0.031,\ 0.013,\ 0.005$ going up to dimension $m = 20$. The convergent curves approach an ε-independent value $K_2(\varepsilon) \simeq 0.3$ in accordance with the estimates in the literature [Grassberger & Proccacia,83b]. Notice that this value is below the Kolmogorov–Sinai entropy $h = 0.42$, which coincides with the largest Liapunov exponent of the system (see section **B.3.3**).

We checked the robustness of the estimate by computing the K_2–entropy for two deformed Henon x–coordinate data series (see figure C.34). Observe that this also exhibits the invariance of the entropy estimates under diffeomorphism. We use the mappings given in figure C.18, and perform the calculation for 30,000 long data series. Notice the consistency of the entropy curves for the dynamics on the synthetic attractors.

The correlation dimension estimate of the logarithms of a x–coordinate series of an orbit of the Henon system obtained in subsection **2.3.2** implies that the K_2–entropy estimate also remains unchanged. In figure C.36 we give the entropy–plot for the series $\{h_1(x_k)\}_{k=1,\ldots,15,000}$, where $h_1(x) = \log(x + 10)$, and $\{x_k\}_k$ is a trajectory of the Henon system. The six upper entropy–curves correspond to radii values from $\varepsilon_1 = 0.1$ to $\varepsilon_6 = 0.0025$, giving right estimates of the true K_2–entropy of the Henon system. We have added the entropy–curve $K_2(0.15)$ (the lowest one in the plot of figure C.36) to illustrate the dependence of the entropy $K_2(\varepsilon)$ on the size ε of the partition considered for its computation. The lower estimate ($\simeq 0.19$) of the entropy given by the curve $K_2(0.15)$ is explained by the compressing effect that the logarithm introduces in the Henon data, and the 'large' size (0.15) of the boxes of that partition. The estimate 0.19 actually gives the degree of stochasticity of the transformed system when observing its dynamics at the resolution scale of this partition. Of course, this estimate gives a proof of the chaoticity of the dynamical system though it does not accurately estimate the intensity of the disorder present in the system. As soon as finer partitions are considered (note that taking $\varepsilon_1 = 0.01$ suffices, which corresponds to the second lowest curve) a more accurate value of the true K_2–entropy is rendered by the entropy curves. Notice that the curve $K_2(0.42)$ of the original Henon attractor does instead give the correct estimation.

We show in figure C.37 the effect on the entropy estimate of corrupting by a small-amplitude noise a series monitored on the Henon attractor. Recall that in **2.2.3** we added what can be taken as experimental noise to an orbit of the Henon system. This is the counterpart of the 'knee' effect explained in section **2.2.3**, and certainly can be predicted from the behaviour of the correlation integrals. The first five entropy curves converge to the right value of $K_2 \simeq 0.3$ but, as soon as the considered partition has the size of the amplitude of the added noise, the entropy curves gives higher values for $K_2(\varepsilon)$. This is caused by the increase of randomness

when the dynamics is scanned within the range where the noise prevails. We address below what the entropy–plot of white noise resembles.

We also computed the K_2–entropies of the Lorenz and the Rössler systems. The length of the x–coordinate series used for the Ruelle–Takens embeddings was 30,000 in both cases. We have plotted in C.32 six entropy–curves for the Lorenz flow (from $\varepsilon_1 \simeq 0.5$ down to $\varepsilon_6 = 0.020$), and eight entropy–curves for the Rössler system (from $\varepsilon_1 = 0.65$ to $\varepsilon_8 = 0.005$) in C.33. In both cases we find fast convergence to positive and finite values of K_2–entropy, which reflect the stochasticity present in both systems. Notice that both values are very close to the estimated values of the positive Liapunov exponents of their respective systems (see section **4.3**).

The entropy–plot of the H_S-series Gi_S is shown in figure C.38. Six entropy–curves are displayed that correspond to radii values from $\varepsilon_1 = 0.22$ to $\varepsilon_6 = 0.09$. The lack of data did not allow us to consider smaller values of ε and, at the same time, to go up to embedding dimension 25. There is a compromise in this analysis between the (smallest) available sizes of the partition and the (largest) value of m that can be used for the estimation in order to keep the analysis significant. It is seen in C.38 that there is no convergence at all in the considered entropy–curves. On the contrary, since the lower curves correspond to the larger values of ε, what the figure shows is the increase of the stochasticity of the system as the size of the partition decreases. This behaviour is compatible with pure–random processes, such as the figure C.39, showing the entropy-plot of an i.i.d. gaussian process, proves. This illustrates the fact that space–filling stochastic processes have divergent K_2–entropy as the size of the partition goes to zero. The random–like behaviour of the K_2–entropy of the H_S–series again supports the rejection of the SAH for our H_S–series.

3.3.2 Correlation Dimension and Entropy Analysis of Filtered SWH–Series. The 'Slow–Cosines' effect.

Analysis of Filtered SWH-Data.

Some significant peaks in the power spectrum can be appreciated in the spectral analysis of the H_S–series. The peaks of the spectrum, when significant, might suggest the hypothesis that some recurrent chaotic component may be active in the H_S–process at some frequency band relevant to the phenomenon. With this idea in mind, we tried to isolate such a chaotic component by filtering the H_S–series, and by applying the correlation analysis to the filtered signals. Since the results obtained from analyzing the series GI_{MO} do not differ from those from the series Gi_S, we concentrate on this latter throughout the rest of this section to illustrate our ideas. Four filtered signals were obtained by low–pass filtering at frequency cutoffs that correspond aproximately to periods below 12 hours (series denoted by $Gif1_S$, cutoff frequencies 0.084 and 0.0875), below one day ($Gif2_S$, 0.0418 and 0.042), below one week ($Gif3_S$, 0.006 and 0.0065), and below 20 days ($Gif4_S$, 0.0021 and 0.00215). In this way, the process is explored through its more significant pseudo–periodic contributions. For the filtering process, we used FFT–routines from I.M.S.L. with a transition of cosine–square type between cutoff frequencies. The filtered series $Gif3_S$ and $Gif4_S$ (their logarithms) are shown in figure C.40.

The correlation plots of series $Gif1_S$ and $Gif2_S$ are given in figure C.42 as well as the $\nu(m)$ vs. m plot. There are no qualitatively significant differences in the behaviour of these two filtered processes compared with the original H_S–process. This belief is enhanced by the analysis of the K_2–entropy, whose computation for the series $Gif1_S$ and $Gif2_S$ display no changes with respect to the entropy of the original signal (which is compatible with pure noise).

The correlation analysis of the signals $Gif3_S$ and $Gif4_S$ seems to be more interesting. It is is shown in figure C.43. The correlation dimension of the series $Gif4_S$ (and also that of $Gif3_S$) seems to saturate at a value aproximately equal to 2 even for the smallest values of m, and the estimate remains stabilized for large values of the embedding dimension. The frequency components remaining in the filtered

series are enough to consider the saturating behaviour of the correlation dimension of $Gif4_S$ in C.43 as non-trivial. The Ruelle-Takens reconstructions of the series in dimensions 2 and 3 (see figure C.41) exhibit a certain 'deterministic' structure, that was not present in the delay plots of the series Gi_S, see figure C.25. Besides, the K_2–entropy of the series $Gif4_S$ (see C.45) does not give a definite indication of pure stochasticity in the reconstructed dynamics of the filtered series; the convergence of the entropy curves to a well–defined positive and finite value rather fits to the SAH. The Liapunov spectrum of the filtered series will be computed in chapter 4, and a test that we have designed to detect chaotic dynamics will be also applied to these series.

Finite Estimates of the Dimension and Entropy for Low–Frequency Stochastic Processes: the "Slow–Cosines" Effect.

The analysis developed so far seems to indicate the presence of a chaotic (non–trivial) component acting in the studied sea evolution over a band of low–frequencies (from periods $\simeq$ 20 days up) masked by the high–frequency component of the physic phenomena. However, such a hypothesis turns out to be wrong. To support this assertion, we created several surrogate series, as described in **A.2.1**, from the series $Gif4_S$. Specifically, this consists of randomizing the Fourier phases of the series $Gif4_S$, and then inverting the spectrum of the signal. In this manner, the phase information of the embedded trajectories is destroyed, so that the correlation integral of an orbit of a true strange attractor of low dimension should behave like that of a purely random process when subjected to this test. This does not turn out to be the case for the filtered series $Gif4_S$. In figure C.44 we show the correlation integral plots of two of these artificial series, which display the same low–dimensional behaviour of the series $Gif4_S$. This makes the hypothesis of the chaotic component be questioned. Actually, we have found this behaviour to be a fundamental characteristic of the correlation dimension estimation of certain stochastic processes with power at low frequencies. In order to show this property, we have built a low–frequency content stationary signal

$$\eta(t) = \sum_{j=1}^{200} A_j \cos(2\pi f_j t + \phi_j), \tag{3.22}$$

by adding up 200 cosines with frequencies $\{f_j\}_j$ equidistributed over the frequency interval where the series $Gif4_S$ is active (that is, the interval $[0, 0.00215]$), and with amplitudes $\{A_j\}_j$ and phases $\{\phi_j\}_j$ uniformly distributed over the interval $(0,1)$ and $(-\pi, \pi)$ respectively. The variance of the series (3.22) was normalized to one. Furthermore, we generated 24,098 data points of the series $\eta(k)$ in order to compare the results with those of the filtered series $Gif4_S$. The delay reconstructions of the series $\{\eta(k)\}_{k=1,2,3,\ldots,23102}$ in dimensions 2 and 3 can be seen in figure C.47, and the correlation integral is shown in C.48. We have also computed the K_2–entropy of (3.22) and the result is given in the top plot of figure C.50. The similarity of these results with those of the signal $Gif4_S$ is obvious. Similar results were obtained for the series $Gif3_S$. The apparent convergence of the correlation dimension estimates $\nu(m)$ to a finite (very low) value close to 2 contradicts the clean stochastic nature of the signal (3.22). This behaviour can be explained by the null–activity of the signal at high frequencies, and the large recurrence times nedeed for the series to properly visit the whole space. The high–frequency components of a noisy process are associated with the filling of the phase space by the embedded orbits, whereas low–frequency noises (like (3.22) is) demand prohibitive data series lengths to reach the same space occupation state. This demand is due to the slowness of the cosines in (3.22), which we thus call *slow–cosines*, in becoming recurrent at a significant level. The results reported above warn against considering finiteness of the correlation dimension as a hint of deterministic chaos.

In order to check the severity of the 'slow–cosines effect' we proceeded in two different ways. We first created a process like (3.22) generated from a sum of 500 cosines whose frequencies ran from 0 to 0.065 (this frequency cut–off allows periodic components larger than 7 days, as the series $Gif3_S$). In spite of this wider frequency interval and the large number of cosines a realization of this process with 50,000 data points was not able to give, either in the correlation dimension or in the entropy analysis, a fair picture of the stochasticity of the process. We show the correlation and entropy plots obtained for this process in figure C.52. Notice that the computations were performed up to dimension 40. Though a very timid increase of the estimated correlation dimensions $\nu(m)$ might be appreciated in the middle plot of figure C.52, observe that $\nu(40) \simeq 2.5$ which is very far of indicating any presence of randomness in the process. The increasing behaviour of the family $\{\nu(m)\}_m$ is expected to improve slightly if larger series of the process are considered, but the

amount of data necessary for a faithful analysis seems unaccesibly large. To check this fact, we secondly performed the computations of the correlation and entropy analysis for a realization of length 50,000 of the process of 200 cosines in (3.22), so duplicating the number of data points with respect to the computations in C.43. We computed the correlation integrals up to dimension $m = 40$, and considered different cut-off delays (see **2.2.2**). The results can be seen in C.51, showing no significant difference with the analysis with 24,098 data points, though a very small increase (the order of $\simeq 0.2$) of the estimate dimension occurs. The entropy plot for this series of 50,000 data points is shown in C.50, giving the same estimate for the K_2-entropy as the series with half its length (see C.50). This confirms the severity of the finiteness of dimension estimates for both the correlation dimension and the K_2-entropy for these kind of stochastic processes.

Notice the slight curvature of the correlation curves in C.49. This gives a hint about the presence of the slow–cosines effect. It is related to the fact that a huge amount of data points will give the correct estimates for a true stochastic process, that is $\nu(m) \simeq m$, at very small resolution scales. What lies at the heart of this spurious effect is the very poor capability of those processes acting at very low frequency ranges to fill (dimensionally) the phase spaces where they are reconstructed. Notice also that the convergence to zero of the K_2–entropy of the series (3.22) (as well as $Gif4_S$) is, to a large extent, related to the finite estimates of the correlation dimension.

The series (3.22) can be regarded as a filtered white noise, and thus with no structure in the power spectrum. This is a difference with the power–law power spectrum noises documented in the literature [Provenzale et al,91], which are shown to give finite estimates of the correlation dimension. We have shown here that no structure is needed in the power spectrum of a stationary signal to yield very low, well–defined correlation dimension estimates of finite–length realizations of pure stochastic processes.

We finish this discussion with some comments on the K_2–entropy estimates of the filtered H_S–series $Gif4_S$, and also of the series of 'slow–cosines'(see (3.22). We show the entropy–plot of the filtered series in figure C.45, in which the entropy–curves

seem to converge to some positive and finite value, suggesting loss of randomness with respect to the primitive series Gi_S. Notice, however, that the entropy–curves are higher as ε decreases and do not intersect each other. The entropies seem thus to increase slightly when ε decreases, and a larger number of data points must yield higher estimates of the K_2–entropies (as occurs for stochastic processes). The situation is the same as that studied above in the correlation dimension analysis. The plot at the top in figure C.50, showing the entropies of a slow–cosines process, implies that finite convergent entropy estimates are also spuriously obtained for finite–length realization of this kind of 'slow' stochastic processes. Of course, the false convergence of the K_2–entropy is related with the stabilizing of the correlation integrals, as can be drawn from equation (3.20). So, it is not surprising that the same convergence problems of the correlation dimension have been also encountered for K_2–entropy [Provenzale et al,91]. Instead of the convergence to zero K_2–entropy claimed to hold for the entropy curves of a family of power–law power spectrum stochastic processes [Provenzale et al,91], we have found a family of white–spectrum stochastic processes with finite and positive K_2–entropy estimate. Of course, this fact warns again about claims of low–dimensional dynamics based just on finite–estimates of the K_2–entropy.

Chapter 4

Testing Liapunov Exponents

There exists a considerable gap in the literature about Liapunov exponents analysis between what has been proved and the cases to which it is being applied in practice. Liapunov exponents are known to exist in the case of continuous and discrete-time flows, where a rigorous theory has been stated. Liapunov exponents analysis is being applied, however, to a wide list of empirical time series, for which there exists no theoretical background. In the introduction below (section **4.1**) we describe the theoretical background of Liapunov exponents, and we explain their significance. In section **4.2** we describe the algorithm that we have designed for the computation of Liapunov exponents. In section **4.3** we comment in detail the various parameters on which the algorithm depends, giving suitable criteria for their selection. Lastly, in section **4.4** we apply the algorithm to the empirical time-series of the significant height of sea waves.

A profound knowledge of the various connections among Liapunov exponents and other dynamic features in the case of smooth flows will be needed for the development of a parallel theory for the stochastic case, where the empirical time series should be modelled. In appendix **B** such connections are shown, sometimes with the help of heuristic arguments and, in any case, with explicit references where the complete proofs can be found. We also describe in the appendix some alternative algorithms to compute Liapunov exponents, and we comment on their differences with ours.

4.1 Introduction.

We first give the so-called *multiplicative ergodic theorem of Oseledec*, key-stone of the theory of Liapunov exponents.

Oseledec's theorem. Let $(M, \mathcal{A}, \mu)$ be a probability space and $f : M \mapsto M$ a measure–preserving transformation. Let $S : M \mapsto \mathcal{M}_{m\times m}$ be a measurable map from M to the real $m \times m$ matrices with

$$\int \log^+ \| S(x) \| \, d\mu(x) < \infty.$$

($\log^+(x) = \max\{0, \log(x)\}$). Write $S_x^{(n)} = S(f^{n-1}) \circ \ldots S(f(x)) \circ S(x)$. Then

a) The limit

$$\Lambda_x = \lim_{n\to\infty} \left(S_x^{(n)*} \circ S_x^{(n)} \right)^{1/2n}$$

exists for μ-almost every x. (We write A^* for the transposed matrix of a matrix $A \in \mathcal{M}_{m\times m}$).

b) Let $\lambda_1(x) > \lambda_2(x) > \ldots > \lambda_k(x)$ be the natural logarithms of the eigenvalues of Λ_x. They are called *characteristic (or Liapunov) exponents* of S at x. (In the case where the smaller eigenvalue of Λ_x is 0, $\lambda_k(x)$ is taken to be $-\infty$). Let $E_i(x)$ be the subspace generated by the eigenvectors corresponding to the eigenvalues $\exp(\lambda_i(x)), \exp(\lambda_{i+1}(x)), \ldots, \exp(\lambda_k(x))$, and let T_x denote the tangent space at x, $T_x = E_1(x) \supseteq E_2(x) \supseteq \ldots \supseteq E_k(x)$. Then, for $v \in E_i(x) - E_{i+1}(x)$

$$\lim_{n\to\infty} \frac{1}{n} \log \| S_x^{(n)}(v) \| = \lambda_i(x). \tag{4.1}$$

c) The functions $x \mapsto \lambda_i(x)$ and $x \mapsto \dim E_i(x)$ are measurable and f-invariant (i.e. they take the same values at $f(x)$ as at x for μ-a.e. x). If μ is ergodic, they are constant μ-a.e.

Remarks.

i) Notice that this theorem does not require any condition on the smoothness either of the measure-preserving map $f : M \mapsto M$ or of the matrix-valued mapping $S : M \mapsto \mathcal{M}_{m\times m}$. This opens the possibility of covering the case of some stochastic processes (see case 4 below).

ii) If the matrices are one-dimensional then $\log \| S(x) \| = \log | S(x) |$, and Oseledec's theorem turns out to be the ergodic theorem of Birkhoff (see **1.4**) for the integrable function $\log |S(x)|$.

iii) If μ is ergodic, the matrix Λ_x need not be constant μ-a.e., although their eigenvalues are, since the eigenspaces of Λ_x usually depend on x.

iv) Concise proofs of this version of the Oseledec theorem can be found in [Ruelle,79] and in [Raghunathan,78]. The theorem was stated by Oseledec [Oseledec,68] under slightly more restrictive assumptions than in the above version, namely the invertibility of the map f.

The standard applications of the theorem are the following:

1. M is a compact smooth manifold, f is a C^r-diffeomorphism, $S(x)$ is the tangent map $D_x f$, and μ is ergodic. The map $x \mapsto D_x f$ is continuous and therefore measurable. The map $x \mapsto \log \| D_x f \|$ is continuous on M compact, so that it is bounded and belongs to $L^1(\mathrm{M}, \mu)$. From the boundedness of $\log \| D_x f \|$ it can also be shown that all characteristic exponents are finite μ-a.e..

Assume that in (4.1) the vector v of the tangent space represents a small perturbation at x that, after n periods of time, evolves to $\| S_x^{(n)}(v) \|$. The (mean) exponential growth of the perturbation over this interval of time is given by the real number λ such that $\| S_x^{(n)}(v) \| = \| v \| \exp(\lambda n)$, so that the limit in (4.1), called *exponent of Liapunov in the direction* v, is just the asymptotic exponential rate of growth of the divergence between the orbit of x and the perturbed orbit of $x + v$.

The matrix $S_x^{(n)}$ may be expressed in polar form (see [Gantmacher,77])

$$S_x^{(n)} = Q_x \circ \left(S_x^{(n)^*} \circ S_x^{(n)} \right)^{1/2},$$

where $Q_x \in \mathcal{M}_{m\times m}$ is an orthonormal matrix and $\left(S_x^{(n)*} \circ S_x^{(n)}\right)^{1/2} \in \mathcal{M}_{m\times m}$ is symmetric. By the orthogonality of Q_x we know that $\| S_x^{(n)}(v) \| = \| \left(S_x^{(n)*} \circ S_x^{(n)}\right)^{1/2}(v) \|$. Therefore

$$\lim_{n\to\infty} \frac{1}{n} \log \| S_x^{(n)}(v) \| = \lim_{n\to\infty} \frac{1}{n} \log \| \left(\left(S_x^{(n)*} \circ S_x^{(n)}\right)^{1/2n} \right)^n (v) \| .$$

A careful analysis of the speed at which the eigenvalues and eigenvectors of the matrix $(S_x^{(n)*} \circ S_x^{(n)})^{1/2n}$ tend to those of Λ_x shows that in the last limit both matrices can be interchanged. Therefore

$$\lim_{n\to\infty} n^{-1} \log \| S_x^{(n)}(v) \| = \lim_{n\to\infty} n^{-1} \log \| \Lambda_x^n(v) \| .$$

This shows that, in the computation of the characteristic exponents, all the factor matrices $S(f^n(x))$ along an orbit can be replaced with the fixed matrix Λ_x, which can therefore be thought of as the *averaged tangent map* along the orbit.

The points x where the Liapunov exponents and the averaged matrix exist are called *regular points*. They are a set of points with full μ-measure. Assume that x is regular. Then the decreasing chain of subspaces in part b) of the Oseledec's theorem completely determines the exponential rate of growth for every vector v of the tangent space. Observe also that, since the dimension of the subspaces is strictly decreasing when i increases, the Liapunov exponent of almost every $v \in E_i(x)$ (with respect to the Lebesgue measure) is λ_i (we may drop the subindex x since the measure μ is ergodic and the Liapunov exponents are constant μ-a.e.). In particular every vector of the tangent space $T_x = E_1(x)$, with the exception of those belonging to the lower dimensional subspace $E_2(x)$ (i.e. almost every vector of T_x with respect to the Lebesgue measure in T_x) has λ_1 as Liapunov exponent. Almost every perturbed orbit diverges at exponential rate from the orbit of x, thus showing the sensitivity of the system to initial conditions (SIC, see **1.1.2**). If, on the contrary, the largest Liapunov exponent is negative, nearby orbits converge at exponential rate, thus denying the possibility of existence of chaotic behaviour. It is for this reason that the computation of the largest Liapunov exponent has become so relevant in the literature.

2. Assume that $f : M \mapsto M$ and $x \mapsto S_x$ are as in **1**, and that the notation stands as there. We shall show how Oseledec's theorem can be applied to the analysis of the

rate of growth of volume elements under the flow. Let $V_x = S_x^{\wedge q}$ be the q-fold exterior product $S_x \wedge S_x \wedge \ldots \wedge S_x$ [Krengel,85]. The mapping $x \mapsto V_x$ satisfies the hypothesis of Oseledec's theorem. Thus, for μ-a.e. $x \in M$ and for a.e. $\{v_1, v_2, \ldots, v_q\}$ (w.r.t. the Lebesgue measure in the q-fold product $\mathbb{R}\times \overset{(k)}{\ldots} \times\mathbb{R}$) the limit

$$\lim_{j\to\infty} j^{-1}\ log\ \| V_x^{(j)}(v_1 \wedge \ldots \wedge v_q)\|, \tag{4.2}$$

with $V_x^{(j)} = V_{fj-1(x)} \circ \ldots V_{f(x)} \circ V_x$, exists.

On the one hand we have

$$\| V_x^{(j)}(v_1 \wedge v_2 \wedge \ldots \wedge v_q)\| = \| S_x^{(j)}(v_1) \wedge S_x^{(j)}(v_2) \wedge \ldots \wedge S_x^{(j)}(v_q)\|;$$

since the right-hand expression is the q-volume of the parallelepiped with edges $S_x^{(j)}(v_1)$, $S_x^{(j)}(v_2), ..., S_x^{(j)}(v_q)$, the limit (4.2) gives the exponential rate of growth of the q-elements of volume under the flow. On the other hand, it can be shown [Krengel,85] that the largest eigenvalue of $A^{\wedge q}$ coincides with the product of the q largest eigenvalues of $(A^* \circ A)^{1/2}$. Thus, if $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_q$ are the q largest Liapunov exponents for μ

$$\lim_{j\to\infty} j^{-1} \log \| V_x^{(j)}(v_1 \wedge v_2 \wedge \ldots \wedge v_q)\| =$$

$$= \lim_{j\to\infty} j^{-1} \log \| S_x^{(j)^{\wedge q}}(v_1 \wedge v_2 \wedge \ldots \wedge v_q)\| =$$

$$= \lim_{j\to\infty} j^{-1} \log \| (S_x^{(j)^*} \circ S_x^{(j)})^{1/2})^{\wedge q}(v_1 \wedge v_2 \wedge \ldots \wedge v_q)\| =$$

$$= \lambda_1 + \lambda_2 + \ldots + \lambda_q,$$

for a.e. $\{v_1, v_2, \ldots, v_q\}$.

3. The Oseledec theorem can be applied for the analysis of continuous–time flows. Let $r \geq 1$ and let F be a C^r vector field in a compact smooth manifold M. Let $f^t(x)$ be the continuous–time flow of the evolution equation $\dot{x} = \mathrm{F}(x)$. Then the time-one mapping f^1 is a diffeomorphism on M, and all the above theory can be adapted to the discrete flow generated by $f \equiv f^1$. This flow coincides, for integer values of t, with f^t. Moreover, a simple continuity argument shows that in all limits ocurring in the Oseledec theorem the discrete time k can be replaced with continuous time t.

4. As we have already remarked, the Oseledec theorem in its full generality does not require the smoothness of f or S, and S is not necessarily any tangent map. The only assumption for the point-to-matrix mapping S is its measurability, and the integrability of $\log^+ \| S \|$. In consequence the theorem can be thought of as an ergodic property of *random products of matrices.* Let us show how this permits the application of Oseledec theorem to the analysis of some stochastic processes. Assume that $\{x_k\}_k$, k integer, is a stochastic process with state space $(N, \mathcal{A}, \nu)$, and let $s : N \mapsto \mathcal{M}_{m\times m}$ be a measurable mapping from the state space to the $m \times m$ matrices. Consider now the infinite-fold product probability space $(M, \mathcal{B}, \mu)$ (i.e. $M = \prod_{-\infty}^{\infty} N$, $\mathcal{B} = \prod_{-\infty}^{\infty} \mathcal{A}$, and $\mu = \prod_{-\infty}^{\infty} \nu$). Then, if the process is stationary, the measure μ is invariant for the forward shift mapping $\sigma\colon M \mapsto M$ which maps the sequence (ω) with i-th coordinate $\pi_i((\omega)) = \omega_i$ onto the sequence $\sigma((\omega))$ with $\pi_i((\omega)) = \omega_{i+1}$. Consider the 0-coordinate mapping $S\colon M \mapsto \mathcal{M}_{m\times m}$ given by

$$S(\ldots\omega_{-1}, \omega_0, \omega_1, \ldots) = s(\omega_0).$$

We can apply the Oseledec theorem taking as f the shift mapping σ, provided that $\log^+ \| S(\omega) \| \in L^1(M, \mu)$. We can also regard the sequence of matrices $(S(\omega_i))_i$ as a matrix-valued stochastic process, and the matrices

$$S_{(\omega)}^{(n)} = s(\omega_{n-1}) \circ s(\omega_{n-2}) \circ \ldots \circ s(\omega_0)$$

are then products of random matrices. In this setting the multiplicative ergodic theorem of Oseledec gives information about the behaviour of such products of random matrices. We shall see below how this allows us to recover the sense of a computation of Liapunov exponents in observed time series by mean of an algorithm based upon estimates of the local linear approximations to the shift mapping.

It should be stressed that there does not exist in the research literature any systematic theoretic development of the application of the Oseledec theorem to the computation of Liapunov exponents in empirically recorded time series *when these are regarded as realizations of stochastic processes.* This does not mean at all that such computation has not been put into practice. On the contrary there exists an extensive literature devoted to the estimation of Liapunov exponents in time series recorded in many research fields, with different algorithms in competence (see references [Brandstäter et al,83], [Wolff,92], [Bajo et al, 92], we turn back to this subject in section **B.4**). Very often the computed Liapunov exponents, when a

positive exponent happens to occur, is taken as evidence of the existence of chaotic dynamics, or at least as a measure of to what extent it is possible to forecast the process. Conclusions of this sort should be managed more carefully, as we shall explain in section **4.4**.

Finally we refer to [Johnson et al,87] for applications of the Oseledec theorem to the analysis of various forms of stochastic differential equations (differential equations with inserted stochastic terms).

4.2 The Algorithm.

The algorithm that we propose for the computation of Liapunov exponents is designed to calculate the exponents along a single orbit, on the implicit assumption that the estimates so obtained approximate to the true exponents of the system by virtue of its ergodicity. In section **4.3.4** we will see how the robustness of such an assumption can be tested.

We assume that the finite orbit $\mathcal{O}_n = \{x_0, x_1, \ldots, x_{n-1}\} \subset \mathbb{R}^m$, from which the Liapunov exponents are computed, is generated either by a discrete dynamical system or by the forward shift of a stochastic process, in such a way that $x_{k+1} = f(x_k)$, $0 \leq k < n$, where f is a diffeomorphism in the deterministic case and the forward shift in the stochastic one.

4.2.1 Linear approximation maps along the orbit.

Once the starting point x_0 of the orbit $\mathcal{O}_n$ has been fixed, the algorithm proceeds to fit the linear mapping β_i which best approximates the forward shift of the orbit, $\sigma(x_i) = x_{i+1}$, about the point x_i. Notice that σ coincides with the restriction of f to $\mathcal{O}_n$, and it is the most natural approximation to f available from the orbit $\mathcal{O}_n$. This step involves in turn two different stages:

- **a) The Search for the neighbours within a fixed distance.** Since we want to obtain local (linear) approximations of f, we only have to take into account the action of the shift in the neighbourhood of the point x_i where we are performing the linear fit, which requires the selection of the points of the orbit within a distance from x_i smaller than a specified ρ. Our algorithm incorporates here the device of the two-dimensional box-assisted technique for an efficient search of neighbours (see chapter **2**). This makes it possible for the algorithm to work along larger orbits and with more neighbours, whilst not consuming too much CPU-time.

- **b) The fit of the linear maps.** This is a process which determines an $m \times m$ matrix S_i for every data point x_i. These matrices play the role of image matrices of the data points under the point-to-matrix mapping in the Oseledec theorem (see **4.1**), i.e., the Liapunov exponents computed from our algorithm are the logarithms of the eigenvalues of the matrix $(S^{(n)^*} \circ S^{(n)})^{1/2n}$, where $S^{(n)} = S_{n-1} \circ \ldots \circ S_0$. In the case of smooth deterministic dynamics, the matrices S_i are intended to be approximations to the tangent maps. In order to obtain such approximations we start by fixing a ρ according to the criteria exposed below (see **4.3.1**). The natural approximation to the ergodic invariant measure from our data is the measure $\sum_{j=0}^{n-1} n^{-1}\delta_{x_j}$, and on a ball B centered at x_i and with radius ρ the relevant measure is $\nu_n = \sum_{j=0}^{n-1} n^{-1} 1_B \delta_{x_j}$, where 1_B stands for the characteristic function of B. The linear approximation S_i is selected as the matrix of the linear mapping L_i which minimizes the functional $\mathcal{A}^p : \mathcal{L} \mapsto \mathbb{R}$ defined on the set $\mathcal{L}$ of linear endomorphisms of $\mathbb{R}^m$ by

$$\mathcal{A}^p(\alpha) = \| \mid \sigma(x_j) - \sigma(x_i) - \alpha(x_i - x_j) \mid_p \|_p,$$

where $\| \cdot \|_p$ is the $L^p(\nu_n)$ norm with $p \in [1, \infty]$ and for $\mathbf{v} = (v_1, v_2, \ldots, v_m) \in \mathbb{R}^m$,

$$\mid \mathbf{v} \mid_p = \left(\sum_{k=1}^{m} \mid v_k \mid^p \right)^{1/p},$$

(in the case of $p = \infty$, $\mid \mathbf{v} \mid_\infty = \max\{\mid v_k \mid : 1 \leq k \leq m\}$).

In other words,

$$\mathcal{A}^p(\alpha) = \left(\int (\mid \sigma(x) - \sigma(x_i) - \alpha(x - x_i) \mid_p)^p d\nu_n(x) \right)^{1/p} = \tag{4.3}$$

$$= \left[n^{-1} \sum_{j=0}^{n-1} 1_B(x_j) \sum_{k=1}^{m} \mid (\sigma_k(x_j) - \sigma_k(x_i) - \alpha_k(x_i - x_j) \mid^p \right]^{1/p},$$

where $\sigma_k(\cdot)$ and $\alpha_k(\cdot)$ are the k-th components of the vectors $\sigma(\cdot)$ and $\alpha(\cdot)$ respectively.

Notice that $\mathcal{A}^p$ attains the same minimum as

$$\sum_{k=1}^{m} \sum_{j=0}^{n-1} 1_B(x_j) \mid \sigma_k(x_j) - \sigma_k(x_i) - \alpha_k(x_j - x_i) \mid^p .$$

This shows that the minimum of $\mathcal{A}^p(\alpha)$ is attained at the linear mapping $L_i = (\alpha_1, \alpha_2, \ldots, \alpha_m)$, where $\alpha_k : \mathbb{R}^m \mapsto \mathbb{R}$, $1 \leq k \leq m$ is the linear form which minimizes

$$\sum_{j=0}^{n-1} 1_B(x_j) \mid \sigma_k(x_j) - \sigma_k(x_i) - \alpha(x_j - x_i) \mid^p .$$

This last problem of minimization can be solved by standard algorithms. In the case where $p = 1$ or $p = \infty$ the uniqueness is not guaranteed. However, in these cases the problem of optimization can be reduced to a problem of linear programming (see [Gonin & Money,89]) for which there exists theoretical background and efficient computational algorithms implemented. For $p \in (1, \infty)$ the optimization of (4.3) is a problem of strictly convex programming with a unique solution. Still there are available standard algorithms to find the numerical solution to this problem.

For $p = 2$ a least squares minimization arises. This is the only case for which an analytical expression for the solution is available. If U is the matrix with the vectors $x_j - x_i$, $x_j \in B$ as rows and V is the matrix with the vectors $\sigma(x_j) - \sigma(x_i)$ as rows, then the matrix of the optimizing linear mapping is given by

$$(\mathrm{U^*U})^{-1}\mathrm{U^*V}.$$

Notice that this expression requires the invertibility of the matrix U*U, whose determinant is the grammian [Gantmacher,77] of the columns of U. If the rows of U are linearly independent the rank of U is m, and the grammian det(U*U) is different from zero. On the contrary, linear dependence of the rows of U entails the non-invertibility of U*U. We return to this subject when discussing the issue of the selection of the optimal value of p (section **4.3.3**).

4.2.2 Computation of the Liapunov exponents. The Q-R algorithm.

Having estimated the matrices S_i, $0 \leq i \leq n-1$, we proceed to the computation of the Liapunov exponents. As we know from section **4.1**, they are given by the limit $\lim_{n\to\infty} n^{-1} \log \| S^{(n)}(v) \|$ for different vectors v of the tangent space. The natural approximation to that limit from our data are the values of the expression

$$n^{-1} \log \| S^{(n)}(v) \|$$

for different values of $v \in \mathbb{R}^m$. We cannot compute such limits because the entries of the matrices $S^{(n)} = S_{n-1} \circ \ldots \circ S_0$ grow exponentially with n and they exceed our available computation power very soon. For this reason we use an algorithm based upon the Q-R decomposition of real matrices, a device for writing the product $S_{n-1} \circ S_{n-2} \circ \ldots \circ S_0$ in an alternative way, which permits the further computation of the Liapunov exponents without performing the product. The method is based upon the following theorem.

> **Theorem (Q-R factorization of matrices).** Given a real matrix $L \in \mathcal{M}_{m\times m}$ there is an orthogonal matrix $Q \in \mathcal{M}_{m\times m}$ and an upper triangular matrix $R \in \mathcal{M}_{m\times m}$ such that $L = QR$. If L is non-singular then R can be chosen positive, and with this choice the factorization is unique.

See [Horn & Johnson,85] for a proof.

There are available efficient numerical algorithms that calculate the Q-R factorization of a matrix. We use the routines LQRR and LQERR of the library I.M.S.L. in our code.

We now show how these results are used to obtain the Liapunov exponents. Let A be a m-dimensional matrix, and assume that $Q_{n-1} \circ R^{(n)}$ is the Q-R factorization of $S^{(n)} \circ A$. From section **4.1** we know that for a.e. matrix $A \in \mathcal{M}_{m\times m}$ (with respect

to the Lebesgue measure in that space)

$$\lim_{n\to\infty} n^{-1} \log \| S^{(n)\wedge q}(Ae_1 \wedge Ae_2 \wedge \ldots \wedge Ae_q) \| =$$

$$= \lim_{n\to\infty} n^{-1} \log \| (S^{(n)} \circ A)^{\wedge q}(e_1 \wedge e_2 \wedge \ldots \wedge e_q) \| = \lambda_1 + \lambda_2 + \ldots + \lambda_q$$

with $e_1 = (1, 0, \ldots, 0) \in \mathbb{R}^m$, $e_2 = (0, 1, 0, \ldots, 0), \ldots$, and $1 \leq q \leq m$. Now it can be readily seen [Krengel,85] that, for any triangular matrix R, $\| R^{\wedge q}(e_1 \wedge e_2 \wedge \ldots \wedge e_q \|$ is just the product of the first q entries of the diagonal of R. Thus

$$\lambda_1 + \lambda_2 + \ldots + \lambda_q = \lim_{n\to\infty} n^{-1} \log \| (Q_{n-1} \circ R^{(n)})^{\wedge q}(e_1 \wedge e_2 \wedge \ldots \wedge e_q) \| =$$

$$= \lim_{n\to\infty} n^{-1} \log \| (R^{(n)})^{\wedge q}(e_1 \wedge e_2 \wedge \ldots \wedge e_q) \| =$$

$$= \lim_{n\to\infty} n^{-1} \log r_{11}^{(n)} r_{22}^{(n)} \ldots r_{qq}^{(n)},$$

where $r_{11}^{(n)} r_{22}^{(n)} \ldots r_{qq}^{(n)}$ are the q first entries of the diagonal of $R^{(n)}$. From this it is readily seen that

$$\lambda_k = \lim_{n\to\infty} n^{-1} \log r_{kk}^{(n)}, \quad 1 \leq k \leq m. \tag{4.4}$$

The algorithm for computing the Liapunov exponents proceeds by performing the Q-R factorization of $S_0 A = Q_0 R_0$, where A is an arbitrary matrix which, for practical purposes, can be set to be the identity matrix. Now we have the product matrix $S^{(n)} \circ A$ written in the form

$$S_{n-1} \circ S_{n-2} \circ \ldots \circ S_1 \circ Q_0 \circ R_0.$$

In the following step the algorithm performs the Q-R factorization of the product $S_1 \circ Q_0$, obtaining matrices Q_1 and R_1 such that $S_1 \circ Q_0 = Q_1 \circ R_1$. Now we can write

$$S^{(n)} \circ A = S_{n-1} \circ S_{n-2} \circ \ldots \circ S_2 \circ Q_1 \circ R_1 \circ R_0.$$

After n steps we arrive at

$$S^{(n)} \circ A = Q_{n-1} \circ R_{n-1} \circ \ldots \circ R_1 \circ R_0.$$

Notice that the product $R^{(n)} = R_{n-1} \circ \ldots \circ R_1 \circ R_0$ is upper triangular. Thus $Q_{n-1} \circ R^{(n)}$ is the Q-R decomposition of $S^{(n)} \circ A$. Since the k-th entry of the diagonal of the product matrix of two upper triangular matrices is the product of

the k-th entries of the factors, if r_{kk}^j denotes the k-th entry of the diagonal of R_j, (4.4) gives

$$\lim_{n\to\infty} n^{-1} \sum_{j=0}^{n-1} \log r_{kk}^j = \lambda_k,$$

and the expressions

$$n^{-1} \sum_{j=0}^{n-1} \log r_{kk}^j, \quad 1 \leq k \leq m$$

are our approximations to the spectrum of Liapunov exponents.

4.3 Significance and Selection of the Parameters and of the Data Form for the Algorithm.

In this section we shall discuss the various parameters which have to be fixed for practical applications of the algorithm, giving some criteria for their selection.

Before proceeding to this discussion, we comment on the results of an adequate selection of the parameters mentioned in the case of the 3–dimensional attractors of the Lorenz and Rössler systems, which will be used as a testing bench throughout this section. See below for the definitons of the parameters. For the Lorenz flow we obtained $\lambda_1 = 1.5063$, $\lambda_2 = -0.007$ and $\lambda_3 = -22.50$ as the Liapunov spectrum, with the input parameters of our algorithm taking the following values: radius $\rho = 0.8$, integration step $\Delta\tau = 0.01$, orbit length $N_{\text{orb}} = 40,000$, number of iterations $N_{\text{it}} = 5,000$, minimum number of neighbours $N_{\text{min}} = 12$ (CPU–time=5.63). For the Rössler flow we obtained $\lambda_1 = 0.0945$, $\lambda_2 = -0.001$ and $\lambda_3 = -9.175$ for the values of the parameters: $\rho = 0.2$, $\Delta\tau = 0.01$, $N_{\text{orb}} = 100,000$, $N_{\text{it}} = 10,000$, $N_{\text{min}} = 23$ (CPU–time=21.32). It should be remarked that finding suitables values for the parameters of the algorithm is not a minor problem. Each data set requires an individual treatment; every parameter for which a choice is possible must be carefully tuned through sucessive proofs, till the required accuracy and robustness in the estimates is achieved. We assume here, as in the previous section, that the orbit $\mathcal{O}_n = \{x_0, x_1, \ldots, x_{n-1}\} \subset \mathbb{R}^m$, from which the Liapunov spectrum is computed, is generated either by a discrete dynamical system or by the forward shift in a stochastic process.

4.3.1 The radius ρ.

This parameter is the radius of the balls $B(x_i, \rho)$ where the linear fit for the shift f is performed (see section **4.2** for definitions). For the case of a differentiable flow f^1, where the tangent maps are known to exist and the linear fits are intended to estimate these tangent maps, too large values of ρ might enhance the weight of the higher order terms in the Taylor expansion of f^1, thus worsening the linear fit. An occular examination of the curvature of the orbits at the scale of the balls $B(x_i, \rho)$ may help to estimate this effect in the case of continuous flows (see figure C.53). It is convenient that the orbits look like small straight line segments at that resolution scale, since the linear fit with positive eigenvalues would decrease the curvature of the orbits if they were bent. In some discrete-time flows of diffeomorphisms the unstable manifold becomes observable (see figure C.1) and it can play a similar role to that of the orbits for continuous-time flows. These sort of considerations are meaningless in the case of stochastic processes where tangent maps are not supposed to exist (see section **4.3.4**).

Large values of ρ could result in an increase of the computing time demanded by the linear fit algorithm. As an example, for the least-squares fit in a ball $B(x_i, \rho)$ with k data points, the number of products of scalars needed is of the order m^2, where m is the dimension of the embedding space.

Small values of the radius ρ improve the linear fit *if they do not result in a lack of data.* This lack of data does not concern only the minimum of m data points in the ball $B(x_i, \rho)$, needed for any linear fit. It concerns in a very significant way the estimates of the negative Liapunov exponents, which could require a number of data points far above this minimum (see section **4.3.3** below for details on this point). See also **B.2.3** for the consequences of the scarcity of data points.

The algorithm described in section **4.2** requires the specification of the p-norm with respect to which the linear fits are performed. Let $S(x_i, \rho, p, n)$ be the matrix of the linear map which minimizes the functional (4.3) with the measure $\nu_n = \sum_{j=0}^{n-1} n^{-1} 1_{B(x_i,\rho)} \delta_{x_j}$ and the norm $\| \cdot \|_p$, with $p \in (1, \infty)$. In [Mera & Morán] it is shown that if f is either a C^1 diffeomorphism, or the time-one map of smooth

continuous-time flow in a compact manifold M and μ is an ergodic invariant measure with respect to f, the limit

$$\lim_{n\to\infty} S(x_i,\rho,p,n) \stackrel{\text{def}}{=} S(x_i,\rho,p)$$

exists for μ-a.e. $x_i \in M$. Moreover, for sufficiently small values of ρ, the map $x \mapsto S(x,\rho,p)$ satifies the hypotheses of the Oseledec theorem. Thus, there exists Liapunov exponents $\lambda_1(\rho,p) \geq \lambda_2(\rho,p) \geq \ldots \lambda_m(\rho,p)$ for these matrices, and they do depend on the value of ρ and on the value of p. Furthermore in [Mera & Morán] it is proved that the limit

$$\lim_{\rho\to 0} \lambda_i(\rho,p)$$

exists and coincides with the i-th Liapunov exponent computed from the tangent maps, independently of p (it can be easily shown that these last do not depend on the norm chosen in (4.1) so that any norm $p \in (1,\infty)$ may be selected to approximate the Liapunov exponents of the tangent maps).

In table C.1 it can be seen the sensitive dependence of the negative exponent of the Lorenz flow on the radius ρ, and the relative stability of the two non–negative exponents with respect to this parameter. Table C.2 affords evidence on the dependence of Liapunov exponents on the radius: when the number of data points increases, thus allowing more accurate estimates on the local fits, the estimates $\lambda_i(\rho,p)$ tend to values different from the true Liapunov exponents (see also comments on table C.15).

4.3.2 Minimum number $N_{\min}$ of ρ-neighbours.

Because of the reasons given above, it might be convenient to fix a minimum of ρ-neighbours of x_i (i.e. data points in each ball $B(x_i,\rho)$) to perform the linear fit. Some authors [Eckmann & Ruelle,85] subordinate the radius ρ to $N_{\min}$ (actually they fix $N_{\min}$ as the exact number of points used to perform the linear fits, see **B.4** for a more detailed analysis of this point): if there are not sufficient ρ-neighbours in the ball $B(x_i,\rho)$ they increase the radius ρ till the neighbours can be found. We proceed in a slightly different manner, keeping fixed ρ while preventing the number of neighbours data points in $B(x_i,\rho)$ to fall below $N_{\min}$. This requires the algorithm

to incorporate a special device for the case in which the ρ-neighbours are less than $N_{\min}$. This is done by searching the least index j larger than i and such that the number of ρ-neighbours of x_j is larger than $N_{\min}$. Then the linear fit of σ restricted to $B(x_{i-1}, \rho)$ is replaced with the linear fit of σ^{j-i+1} in the same domain. The value of $N_{\min}$ should not be maintained as large as to cause the difference j-i+1 to be too large for some i. The described method has the advantage of keeping ρ under control, which gives a theoretic support to the Liapunov exponents computed with finite data sets as we have seen above. This is specially convenient in the case of stochastic processes, where the Liapunov exponents computed for decreasing values of ρ might not be convergent (see section **4.4**).

4.3.3 Selection of the p–norm.

We now show how the dependence of the exponents $\lambda_i(\rho, p)$ on p might be used, at least in some cases, to improve the estimates of negative Liapunov exponents from finite data sets. In **B.2.2** and **B.2.3** it is argued that for SRB ergodic measures and in some other cases, as the Lorenz attractor, the invariant measure is expected to behave as a measure absolutely continuous w.r.t. the Lebesgue measure on the unstable manifold and as a singular measure along the stable manifold. Once the value of ρ is fixed in the largest range possible as to keep under control the non–linearity of the flow, it might happen that the number of ρ-neighbours in the ball $B(x_i, \rho)$ is still low. In this case it is expected that most of the data points are lying in the neighbourhood of the linear subspace tangent to the unstable manifold, which is generated by the eigenvectors of the matrices Λ_{x_i} corresponding to the positive Liapunov exponents (see **B.1** and **B.2.2**). Since the algorithms implemented to estimate Liapunov exponents compute averaged rates of growth of the vectors of the tangent space *along the directions provided by the data points*, we can only expect satisfactory estimates of the positive exponents (see table C.1). When using the 2-norm with scarcity of data points, the fact that these are lying approximately in the unstable manifold can cause the quasy-linear dependence of the rows of the matrix U and therefore the matrix $(U^*U)^{-1}$ needed for the fit (see **4.2.1**) to be badly determined. For this reason we incorporate the same device as in **4.3.2** to reject the points x_i where this happens. This can be done by fixing a suitable parameter

in the subroutine of regression of the IMSL.

On the other hand, we have experimental evidence that the selection of the 1-norm and ∞-norm significantly improves the estimates of the negative Liapunov exponents in a situation of lack of data (for the Lorenz flow we obtained $\lambda_3 = -19.02$ using 2–norm and $\lambda_3 = -22.12$ using 1–norm, for algorithm parameters $\rho = 0.8$, $N_{\rm orb} = 40,000$, $N_{\rm it} = 1,000$, and $N_{\rm min} = 6$). Although for these norms there does not exist any result such as that described at the beginning of this section, they are also suitable to make the linear fits which turn out to be for both norms a problem of linear programming (on the contrary for all other values of p, which lead to problems of non-linear convex programming) with efficient algorithms available.

We now discuss separately two cases where the algorithm can be used for the computation of the Liapunov spectrum, namely chaotic attractors generated from known evolution equations, and empirically observed stationary time series under the *strange attractor hypothesis* (SAH). The case of empirically observed stationary time series generated by a stochastic process is analyzed in section **4.4**.

4.3.4 Data form for chaotic attractors with given equations: integration step $\Delta\tau$, orbit length $N_{\rm orb}$ and number of iterations $N_{\rm it}$.

The evolution equations generating the chaotic motion give the *embedding dimension* m, i.e. the dimension of the space where the attractor lives. This parameter coincides with the cardinality of the set of variables characterizing the state of the system and of the set of equations. In this case we generate our data as a series $x_0, x_1, \ldots x_{n-1}$ of m-dimensional vectors to which we apply the algorithm.

An evolution equation of the form $\dot{x} = F(x)$ (continuous-time flow) is either trivial or non-integrable. Explicit expressions for the flow f^t cannot therefore be obtained, and we need linear fits to get estimates of the tangent maps. We have to generate to this end an orbit from which, by virtue of the ergodicity of the system, we can draw the Liapunov exponents. It is advisable to generate that orbit with some

accurate algorithm of numerical integration, namely a fourth or sixth order Runge-Kutta scheme. There are available routines in the I.M.S.L. The sketched scheme should be implemented with a step of integration $\Delta\tau$ short enough as to ensure as much accuracy as possible in the numerical integration of the orbits. This step depends on the "speed" of the dynamics. The numerically integrated orbits should appear like smooth curves rather than chains of linked straight line segments. In the case of Lorenz or Rössler attractors, the recommended step is of the order of 10^{-3}.

The integration time step is usually shorter than the *sampling time* Δt. The sampling time is then some multiple of $\Delta\tau$ (for the Lorenz and Rössler attractors, of the order 0.03) determined by the speed of the motion. It should be taken sufficiently small as to ensure that the flow $f^{\Delta t}$ can be well approximated by the linear fits. The average spatial change $f^{\Delta t}(x) - x$ of a point x of the attractor should ensure some sameness between the neighbourhoods of the attractor around both points, x and $f^{\Delta t}(x)$. By virtue of the smoothness of the vector field F this is guaranteed if Δt is small enough. Too small values of Δt have, however, two negative effects: first, they entail an increase of the number N_{it} of data points needed to visit the whole attractor, which could cause too much waste of CPU-time for the computation of the Liapunov exponents (see below comments on the number N_{it} of iterations); second, they may worsen the linear fits, specially in the direction of the stable manifold when the fits are obtained with few data points. In any case, a *plateau* where the exponents are stabilized with respect to the sampling time should be found if the exponents of a chaotic dynamics are robust (see table C.4 for a test on the Lorenz attractor).

In numerically simulated dynamics, the number N_{orb} of data points has to be chosen large enough to ensure that the whole attractor is visited by the orbit several times. A guideline for determining this parameter is the following: if ρ has been fixed according to the criteria set out above (section **4.3.1**), then N_{orb} has to be large enough to ensure more than N_{min} visits to the balls $B(x_i, \rho)$ (some but not too many exceptions are allowed, see **4.3.2**). Any increase of N_{orb} would permit a decrease of ρ, thus improving the estimates of the tangent maps and therefore the Liapunov exponents, especially the positive ones. If we maintain ρ fixed an increase of N_{orb} would improve the estimates of the negative Liapunov exponents over balls

of small radii, as shown in table C.2.

For discrete-time flows with equations of the form $x_{k+1} = f(x_k)$ $k \in \mathbb{N}$, $x_k \in \mathbb{R}^m$, and $f \in C^1(\mathbb{R}^m)$, the matrix S_i of the tangent map at x_i can also be obtained from the equations as the Jacobian matrix of f at x_i, which can be drawn from the equations. Thus we can ignore the step of linear fits and apply the Q-R algorithm to these matrices (see **4.2.2**). There does not exist here a step of integration, the sampling time is one, and the criteria for the selection of N_{orb} are as above.

The parameter N_{orb}, if selected according to the above criteria, is usually unsuitably large to coincide with the number N_{it} of matrices in the sequence $\{S_i\}_{1 \le i \le N_{\text{it}}}$ to which we apply the Q-R algorithm (see **4.2.2**). The formulae (4.4) may converge for values of n far smaller than N_{orb}, as suggested in table C.3. Since the Liapunov exponents are averaged rates of divergence between nearby orbits, their computation requires a sampling of nearby orbits throughout the whole attractor. From this point of view the parameter N_{it} should be so large as to ensure that the sequence of balls $\{B(x_i, \rho)\}_{1 \le i \le N_{\text{it}}}$ covers the entire attractor (compare with the criterion given above for N_{orb}). For the attractors of our testing laboratory suitable values are 500 (Henon), 2,000 (Lorenz) and 10,000 (Rössler).

A useful test of the robustness of the Liapunov exponents estimates, once the parameters have been chosen, is their invariance with respect to the starting point, according to the fact that the ergodicity of the invariant measure implies that the estimated Liapunov exponents are the same for μ-almost every starting point. A right selection of parameters should result in small variances of the distribution of the Liapunov exponents $\lambda_i(x_0)$ computed from the chosen parameters with starting point x_0. See table C.5 for a test on the Lorenz attractor.

4.3.5 Data form for empirical time series under the SAH: the embedding dimension E, and the linear fitting dimension F.

We now want to analyze a univariate empirically recorded time series $u_0, u_1, \ldots u_n$, $u_i \in \mathbb{R}$, $1 \leq i \leq n$. We first select an *embedding dimension* E, then we vectorize our data as a sequence of E-histories $\{x_i^{(E)}\}_{0 \leq i \leq n-E+1}$ (recall that $x_i^{(E)} = (u_i, u_{i+1}, \ldots u_{i+E-1})$), and lastly we compute the Liapunov exponents from these data. We now discuss what is the meaning of this computation. We assume that the *strange attractor hypothesis* (SAH) holds (see section **1.1.2**). This means that there exists a compact manifold $M \subset \mathbb{R}^M$ and an evolution equation $y_{k+1} = f(y_k)$ where $f : M \mapsto M$ is the time-one mapping of a smooth deterministic flow, and that the observed series $u_0, u_1, \ldots u_n$ has been obtained by a scalar projection h from $\{y_i\}_{i \in \mathbb{N}}$, i.e. $u_i = h(y_i)$, $h : M \mapsto \mathbb{R}$ being a smooth real function. By virtue of the theorem of Takens (see section **1.3**), if $E \geq 2M + 1$ the mapping $J : M \mapsto \mathbb{R}^E$, given by

$$J(y) = (h(y), h(f(y)), \ldots h(f^{E-1}(y))),$$

is generically a diffeomorphism onto its image $J(M)$. If $g = J \circ f \circ J^{-1}$,

$$g(x_i^{(E)}) = g(u_i, u_{i+1}, \ldots, u_{i+E-1}) =$$

$$= J \circ f \circ J^{-1}(h(y_i), h(f(y_i)), \ldots, h(f^{E-1}(y_i))) = J \circ f(y_i)$$

$$= J(y_{i+1}) = (h(y_{i+1}), h(f(y_{i+2})), \ldots, h(f^{E-1}(y_{i+1}))) = x_{i+1}^{(E)}$$

Therefore $\{x_i^{(E)}\}_{i \in \mathbb{N}}$ is an orbit of the discrete flow $\{g^k\}_k$ generated by the diffeomorphism $g = J \circ f \circ J^{-1}$ on the compact manifold $J(M) \subset \mathbb{R}^E$, and the Liapunov exponents computed from the finite sequence $\{x_i^{(E)}\}_{0 \leq i \leq n-E+1}$ are those of such a flow. The diffeomorphism J maps the M-dimensional tangent space T_y at every $y \in M$ onto a M–dimensional subspace $J(T_y)$ of the E-dimensional tangent space $T_{J(y)}$. Let $x \in J(M)$ and $y = J^{-1}(x) \in M$,

$$S_x^{(k)} = D_{g^{k-1}(x)} g \circ D_{g^{k-2}(x)} g \circ \ldots D_x g,$$

and

$$R_y^{(k)} = D_{f^{k-1}(y)} f \circ D_{f^{k-2}(y)} f \circ \ldots \circ D_y f.$$

Since $D_{g^i(x)}g = D_{f^{i+1}(y)}J \circ D_{f^i(y)}f \circ D_{g^i(x)}J^{-1}$, for vectors $v \in J(T_x)$ we have

$$\lim_{k\to\infty} k^{-1} \log \| S_x^{(k)}(v) \| =$$

$$= \lim_{k\to\infty} k^{-1} \log \| D_{f^k(y)}J \circ R_y^{(k)}(D_x J^{-1}(v)) \| .$$

From the compactness of M follows the existence of constants K_1 and K_2 such that $K_1 \| w \| \leq \| D_x J(w) \| \leq K_2 \| w \|$ for any $x \in M$ and $w \in T_x$. This shows that

$$\lim_{k\to\infty} k^{-1} \log \| S_x^{(k)}(v) \| = \lim_{k\to\infty} k^{-1} \log \| R_y^{(k)}(D_x J^{-1}(v)) \| .$$

Notice that the right-hand term is the Liapunov exponent in the direction $D_x J^{-1}(v)$ for the flow $\{f^k\}_k$, whilst the left-hand term is the Liapunov exponent in the direction v for the flow $\{g^k\}_k$. This shows that the Liapunov exponents of $\{g^k\}_k$ in the directions of the subspaces $J(T_y)$ for $y \in M$ are those of $\{f^k\}_k$. Observe that, in particular, the invariance of the Liapunov spectrum under diffeomorphisms follows from above.

The E-M Liapunov exponents of $\{g^k\}_k$ in directions other than those of the subspaces $J(T_y)$ are called in the literature *spurious Liapunov exponents.* When computing the Liapunov exponents of $\{f^k\}_k$ from the embedded sequence $\{x_i^{(E)}\}_{0\leq i\leq n-E+1}$ we have no *a priori* hint of what exponents are spurious. We refer to [Abarbanel et al,91] where a method is proposed to distinguish positive spurious exponents. We give below a criterion which could be useful for both positive and negative exponents.

If the SAH holds, the embedded sequence $\{x_i^{(E)}\}_{0\leq i\leq n-E+1}$ of E-dimensional vectors is expected to be lying in the neighbourhood of the M-dimensional subspaces $J(T_y)$. This means that the matrices U and V used to perform the least squares fits could not be of full rank. The same problem is also likely to influence negatively the linear fits under selections of the p-norm other than two. For this reason, Eckmann, Kamphorst, Ruelle and Ciliberto [Eckmann et al,86] have proposed the projection of the vectors $\{x_i^{(E)}\}_{0\leq i\leq n-E+1}$ onto lower dimensional subspaces. Let F be any integer such that E-1 is a multiple of F, and let π be the projection $\pi : \mathbb{R}^E \mapsto \mathbb{R}^F$ defined by

$$\pi(u_i, u_{i+1}, \ldots, u_{i+E-1}) = (u_i, u_{i+p}, u_{i+2p}, \ldots, u_{i+Fp}),$$

with $Fp = E - 1$. Since $\pi(E)$ is an F-dimensional space, for $r \leq F$, almost every r-dimensional linear subspace (with respect to the natural measure in the Grassmanian manifold G(r,E) (see [Mañé,81]) of the r-dimensional linear subspaces of $\mathbb{R}^E$

projects under π onto an r-dimensional subspace. Except for those $y \in M$ such that $\dim(\pi(J(T_y))) < F$, $H = \pi \circ J$ is a submersion (see [Guillemin & Pollac,74]). We can perform a linear fit between the F-dimensional projections of ρ-neighbours of $x_i^{(E)}$ and the F-dimensional projections of their images under g. In particular, if F= M, the surjectivity of π guarantees that such fits are not badly determined, and in general any value of F lower than E improves the problem of bad determination. Since H could not be injective we cannot follow the above reasoning to show that the Q-R algorithm applied to such linear fits gives the Liapunov exponents for the flow f. The quoted authors claim that this should be expected to occur, since we do not entangle the dynamics with this submersion in a lower dimensional subspace, as should occur if we would make the fits with ρ-neighbours in the projected space (which could not be neighbours at all in $\mathbb{R}^E$) instead of considering projections of ρ-neighbours in $\mathbb{R}^E$. An interesting consequence of the correctness of this method would be the fact that under different choices of F the true Liapunov exponents of f would always appear, thus providing a criterion to detect the spurious exponents. For these reasons we have included the fitting dimension as a parameter in our code.

4.4 Liapunov Exponents computed from Empirically Recorded Time Series generated by a Stochastic Process: Application to Significant Wave Height Time–Series.

4.4.1 General Remarks.

We now assume that our recorded data $u_0, u_1, \ldots, u_N$ are a sample of a stochastic process $\{X_i\}_{i\in\mathbb{N}}$, where $X_i : \Omega \mapsto \mathbb{R}$ is a random variable on a probability space $(\Omega, \mathcal{A}, p)$ and $u_i = X_i(\omega)$ for some $\omega \in \Omega$ and for every $i \in \mathbb{N}$. In many cases this stochastic process is drawn from another one of the form $\{X_t\}_{t\geq 0}$ by setting $X_i = X_{i\Delta t}$. The parameter Δt plays here the same role as the *sampling time* in section **4.3** above. A shortening of the sampling time does not necessarily improves the data quality. This depends on the smoothness of the sample functions $X_t(\omega)$ when viewed as functions of t. A reduction of Δt which does not result in an increase

of the maximum slope in the u_i vs. i plot could be useless, and, in fact, it could be impractical because of the artificial enlarging of the data set. An enlarging of the sampling time interval $n\,\Delta t$ will often improve the data quality, giving a more complete account of the long–term behaviour of the process and allowing more accurate estimates of the parameters of the process.

In our new setting the relevant questions are the following: does the computation of Liapunov exponents from these data make sense?, and what do such exponents tell us about the process?.

We proceed to formulate some tentative answers to the above questions. The convenience of starting by imposing the hypothesis of stationarity of the process is apparent. Let us recall the meaning of this assumption. The stochastic process defines a probability space $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}^{\mathbb{N}}, \mu)$, where μ is a measure defined on the σ–algebra $\mathcal{B}^{\mathbb{N}}$ of subsets of $\mathbb{R}^{\mathbb{N}}$ generated by the cylinder sets of the form

$$\{(y_1, y_2, \ldots) \in \mathbb{R}^{\mathbb{N}} : \ (y_j, y_{j+1}, \ldots, y_{j+p-1}) \in A, \ A \text{ is a Borel subset of } \mathbb{R}^p, \ j, p \in \mathbb{N}\}.$$

To say that the process is stationary is the same as to say that the measure μ is invariant for the *forward shift mapping* $\sigma : \mathbb{R}^{\mathbb{N}} \mapsto \mathbb{R}^{\mathbb{N}}$, $\sigma(y_1, y_2, \ldots) = (y_2, y_3, \ldots)$. Under the further assumption of ergodicity of the process, for any integrable function $f : \mathbb{R}^{\mathbb{N}} \mapsto \mathbb{R}^{\mathbb{N}}$ we have

$$\int f d\mu = \lim_{k \to \infty} k^{-1} \left(f(\mathbf{y}) + f(\sigma(\mathbf{y})) + \ldots + f(\sigma^{k-1}(\mathbf{y})) \right)$$

for $\mathbf{y}$ belonging to a subset of $\mathbb{R}^{\mathbb{N}}$ with μ–probability one. This allows us to estimate a given parameter of the process (as the mean, the variance, moments of higher orders, etc.) by averaging the given parameter throughout a single sample of the process. If a unique sample of our time series is available, the ergodicity of the process is needed in order to extrapolate the observed properties of the series to the whole process. Milder conditions of partial ergodicity on the process, i.e. ergodicity with respect to particular parameters can also be considered. This involves the assumption that the equality above holds for such parameters, but it can fail for other integrable functions.

Let $f : \mathbb{R}^{\mathbb{N}} \mapsto \mathbb{R}^m$ be any integrable function with respect to the joint distribution $p_{(m)} = p_{12\ldots m}$ of the random variables $X_1, X_2, \ldots, X_m$, and let $\pi_m : \mathbb{R}^{\mathbb{N}} \mapsto \mathbb{R}$

be the projection given by $\pi_m(y_1, y_2, \ldots) = (y_1, y_2, \ldots, y_m)$. If the process $\{X_i\}_{i\in\mathbb{N}}$ is ergodic

$$\int f dp_{(m)} = \int f \circ \pi_m d\mu =$$

$$\lim_{k\to\infty} k^{-1} \left(f \circ \pi_m(\mathrm{y}) + f \circ \pi_m(\sigma(\mathrm{y})) + \ldots + f \circ \pi_m(\sigma^{k-1}(\mathrm{y})) \right) =$$

$$\lim_{k\to\infty} k^{-1} \left(f(x_1) + f(x_2) + \ldots + f(x_k) \right),$$

where we have written x_j for the m–history $\pi_m(\sigma^{j-1}(\mathrm{y})) = (y_j, y_{j+1}, \ldots, y_{j+m-1})$. Thus the integration of f can be performed by averaging the values of f over the m–histories. In particular, if we take as f the indicator 1_B of a Borel set B, the asymptotic rate of times that the m–histories belong to B gives the $p_{(m)}$–measure of B. This shows that algorithms such as that of Grassberger–Procaccia for the computation of the correlation dimension described in chapter **2** make sense for ergodic stochastic processes, since they give us estimates of the dimension of the joint probability distribution of consecutive variables of the process. We now apply the Eckmann–Ruelle algorithm for the computation of Liapunov exponents to SWH–series described in section **2.3.1**, and show how this algorithm can be used as a test for checking the SAH.

4.4.2 Application to SWH–series.

In tables C.6, C.9, and C.10 we can see the results of applying our algorithm for the computation of Liapunov exponents to the series Gi_S described in section **2.3.1**. We performed the computations over intervals of time for which the record either is complete or the intermittences of missing data are very short. The algorithm is applied to reconstructions of the series in dimensions one, two and three, performing the linear fits over several values of the radius ranging between $\rho = 0.05$ and $\rho = 0.5$ (see tables C.6, C.9, C.10). We see that, despite the obvious randomness of the underlying stochastic process, the Liapunov exponents are always very near zero, having even slightly negative values in some cases. This shows that the averaged linear fit is near the identity.

There is a remarkable independence of the computed Liapunov exponents from the starting point of the orbit, which proves the existence of ergodic properties of the

process with respect to parameters such as the Liapunov exponents which are not of simple definition, and which are computed through the complex algorithm described in the previous sections of this chapter. This can be seen in the tables C.6, C.7, and C.10, C.11, where the same Liapunov exponents as above are computed for different starting points. Tables C.6 and C.8 show that the computed Liapunov exponents remain stabilized under further increases of the number of iterations (parameter N_{it} of the algorithm, see section **4.3.4**).

Observe in table C.10 that only in dimension three and when the radius of the balls becomes very small does the maximum Liapunov exponent increase up to significantly positive values. Does this mean that there is no sensitivity to initial conditions for nearby orbits within distances in the range corresponding to the larger values of ρ giving Liapunov exponents close to zero or negative?. Should we assert that the time series is neither stochastic nor even chaotic?. We will try to explain the above results from an alternative point of view, supporting our theoretical explanation on empirical evidence obtained from other chaotic and stochastic processes.

The forward shift mapping induces a forward shift $\sigma(x_i) = x_{i+1}$ in the set of m–histories. If we use these m–histories as input for the Eckmann–Ruelle algorithm, we are computing the best linear fit (according to some chosen p–norm criterion) to the 'action' of the forward shift on the set of m–histories. This provides a sequence of stochastic matrices for which the Oseledec theorem, under suitable hypothesis of measurability and integrability of the relevant maps, could still guarantee the existence of a Liapunov spectrum depending on the p–norm and on the radius ρ of the balls selected.

Unfortunately the action of the shift is not that of a differentiable mapping. In fact, it need not be even a mapping: it may occur that $(u_i, u_{i+1}, \ldots, u_{i+m-1}) = (u_j, u_{j+1}, \ldots, u_{j+m-1})$ but $u_{i+m} \neq u_{j+m}$, and this would imply that $x_i = x_j$ but $\sigma(x_i) \neq \sigma(x_j)$. In this case it is clear that the Liapunov exponents as they are usually understood do not make sense.

4.4.3 A New Test for Detection of Chaos in Time–Series.

We claim that the Eckmann–Ruelle algorithm can be used, however, as a test for testing the SAH. This test can be applied to any time series, although it seems specially suitable for time series that do not reject the SAH from the Grassberger–Procaccia dimensional test, that is, for time series whose m–dimensional Ruelle–Takens reconstructions do not increase the estimates of their correlation dimension –obtained from the Grassberger–Procaccia algorithm– when m increases. The slow–cosines effect (see section **3.3.2**) shows that further analysis is needed to ensure the existence of chaotic dynamics. With that aim we propose the computation of a parameter $r_i(\rho,p)$ (see below for a definiton) which estimates the correctness of the best L^p–linear fit to the forward shift mapping σ in a ball of radius ρ about the i-th m–history x_i. If the process is the projection on $\mathbb{R}^n$ of an ergodic system, the value

$$r(\rho,p) = N^{-1}\sum_{i=1}^{N} r_i(\rho,p)$$

of the parameter $r_i(\rho,p)$ averaged throughout the orbit gives information about the mean rightness of such linear fits as approximations to the forward shift σ. If the data are generated by a dynamical system, the parameter should range in a certain interval close to one. Furthermore, a reduction of the radius ρ which does not result in a reduction of the number of data points in the balls of radius ρ centered at the data points, should increase the correctness of the fits, and therefore it should reduce the value of $r(\rho,p)$. These could be considered strong indications of the existence of chaotic dynamics, although one cannot still guarantee the existence of such dynamics, since some stochastic processes might be well approximated by local linear fits. If, on the contrary, the values of $r(\rho,p)$ remains far from one, the SAH can be definitely ruled out, at least for the tested embedding dimensions.

Let m be fixed. The parameter we propose to choose is

$$r_i(\rho,p) = \min_{1\le j\le m}\{r_i(j,\rho,p)\},$$

$r_i(j,\rho,p)$ being given by, for $1 \le j \le m$,

$$r_i(j,\rho,p) = 1 - \frac{\sum_{k=1}^{N_{obs}(\rho,i)} \mid y_{k+1}^j - \hat{y}_{k+1}^j \mid^p}{\sum_{k=1}^{N_{obs}(\rho,i)} \mid y_{k+1}^j \mid^p},$$

where y^j_{k+1} is the j-th coordinate of the vector $y_{k+1} = x_{k+1} - x_{i+1}$, $x_k \in B(x_i, \rho)$, $\hat{y}^j_{k+1}$ is the j-th coordinate of the vector image of y_k under the fitted linear mapping, and $N_{\text{obs}}(\rho, i)$ is the number of m–histories in the ball $B(x_i, \rho)$. This parameter takes values in the interval $[0, 1]$. A value of $r_i(\rho, p)$ near *one* indicates a good linear fit, while a value $r_i(\rho, p)$ near *zero* indicates a bad fit. In tables C.14 and C.15 the averaged values $r(\rho, 2)$ for the Henon map and for the Lorenz system for different values of ρ can be seen. As it was expected to be, the values of $r(\rho, 2)$ are close to one in both cases, and they increase when ρ tends to zero. This supports the conjecture that if the SAH holds, the parameter should take high values, tending to one when ρ tends to zero. Theoretical work is in progress [Mera & Morán] supporting this conjecture.

In tables C.16 and C.17 the parameter $r(\rho, 2)$ is computed from the one, and two–dimensional reconstructions obtained from an i.i.d. of random variables with uniform distribution. It takes values near from 0, thus indicating bad linear fits. Figure C.54 gives account of the existence of a power–law dependence between the radius of the ball and the Liapunov exponent in the one–dimensional case. We think that there is a statistical compensation of the spatial directions of spreading, due to the randomness of the process, which causes the best linear fit to be near the identity; or even negative for large values of the radius ρ. Assume for simplicity that we are working with a one–dimensional series at a point u_i. Assume that the neighbours of u_i laying on an interval with radius ρ centered at u_i are spread out by the forward shift over an interval of values with length $2\rho'$ with $\rho' \gg \rho$. However, if the neighbours on both half intervals are scattered by σ uniformly over the interval $(u_i - \rho, u_i + \rho)$, the identity map could still be a reasonable fit for σ, even though in the case of the uniform distribution we know, from the computation of the parameter $r(\rho, 2)$, that such fits are rather deficient. This also explains, from our point of view, the values close to zero for the Liapunov exponents displayed in table C.6.

Observe also that when the radius ρ of the balls decreases, the maximum Liapunov exponent increases, becoming positive (see tables C.10 and C.11). This could be due to the fact that when there are scarcity of data points in the balls considered, the above mentioned statistical compensation cannot occur; the algorithm is determined by a few close neighbours to every point, the linear fits become more accurate (this is consistent with the value of $r(\rho, 2)$ obtained in table C.17), and the

divergence of their orbits can be detected by the algorithm.

The computation of the values of the parameter $r(\rho, 2)$ from the SWH–series are shown from table C.6 to table C.12. They prove that the averaged best linear fit for the shift σ is a bad one, and that it worsens when ρ decreases. This should be interpreted, from our point of view, as showing that the SAH does not hold in these series, at least for the embedding dimensions one to ten (see table C.12). The test cannot properly be applied to higher embedding dimensions because of the scarcity of data. Observe in table C.12 that when the dimension increases the parameter $r(\rho, 2)$ improves. This is because there are more available parameters to perform the linear fits. However, an increase of the number of N_{orb} of data points could worsen the value of $r(\rho, 2)$, which in any case is always far from the values obtained for the chaotic dynamics. In the tables in C.13 it is displayed the dependence of the parameter $r(\rho, 2)$ on the sampling time for the Gi_S–series.

In tables C.18, C.19, C.20, C.21 we test the Liapunov exponents of the filtered $Gif4_S$ series. Due to the low frequency content of the process, the time flow for low sampling times is near the identity and the Liapunov exponents are near zero. An increase of the sampling time up to the value 50, which only allows us 20 iterates (provided that the series gaps are to be avoided), causes the value of some Liapunov exponents to become non–null and an abrupt decrease of the parameter $r(\rho, 2)$: compare the first columns of tables C.18 and C.20. The latter is close to one for embedding dimensions larger than two. Tables C.19 and C.21 show, however, that the qualitative behaviour of the Liapunov exponents of the $Gif4_S$ series is quite similar to that of the $Gif4_S$ phase–randomized series, and that of the slow–cosines processes (see tables C.19 and C.20 respectively).

w

4.4.4 An Algorithm to obtain Estimates of the Divergence between Nearby Orbits in Stochastic Processes. Application to SWH–series.

According to the previous paragraph, the Eckmann–Ruelle algorithm could not make sense in the case of stochastic systems. Other algorithms reviewed in this work (see **B.4**) make no sense either, since they require the asymptotic properties of alignement with the direction of the unstable manifold, which need not exist for an ergodic stochastic process. A more direct approach for the computation of the divergence of nearby orbits is proposed by Wolff [Wolff,92] who performs a direct computation of the divergence of pairs of orbits about a given point, and averages the rates so obtained. This computation has significance on its own accord, although following the above analysis it does not give the largest Liapunov exponent, despite the title given by its author to the article in which these ideas are developed.

We proposed the computation of another parameter, the *rate of norm–divergence*, which supports the arguments given in the previous paragraph and give an alternative approach for the estimate of the divergence of nearby orbits of a stochastic process. This parameter approximates to the largest Liapunov exponent in time series recorded from differentiable dynamical systems. It avoids the problem remarked above when applied to stochastic systems, thus providing an estimate of the sensitivity to initial conditions of such systems. We propose to modify the Eckmann–Ruelle algorithm so that it computes at each point x_i the best linear fit to the one–dimensional mapping that maps the L_p–norm $\| x_i - x_j \|_p$ of the vectors of the tangent space with $x_j \in B(x_i, \rho)$ onto the L_p–norm $\| x_{j+1} - x_{i+1} \|_p$ of their images. This gives a linear approximation to the rate of growth of the L_p–norm of a perturbation about the point x_i of the orbit. We then compute the unique Liapunov exponent obtained from these linear fits throughout the orbit. Since we are ignoring the spatial direction of spreading of nearby points, we are eliminating the statistical compensation mentioned above, and the Liapunov exponent computed in this way should recover the real significance of the largest Liapunov exponent as the averaged exponential rate of growth of the divergence of nearby trajectories. This can be confirmed in the tables in C.25 performed over the Henon and Lorenz attractor, where values near the true largest Liapunov exponents are obtained. See also table C.26,

and compare it with table C.16, where the conventional Liapunov exponents for a uniformly distributed independent sequence of random variables was computed. In figure C.55 it is displayed the power–law dependence of the orbital norm–divergence on the radius.

In table C.22 we can see the estimates of the rate of norm–divergence for the 3–histories of the Gi_S series. The values of the Liapunov exponents become positive (compare this with table C.6) and increases when ρ decreases. The values of the parameter $r(\rho, 2)$ are higher than those in table C.6, especially for low values of ρ. The computation is extended for dimensions up to 10 in table C.23. This can be compared with table C.12, thus verifying that the same remarks made above for dimension three remain true. In table C.24 it is checked the dependence of the orbital norm–divergence on the sampling time and on the number of iterations.

Chapter 5

Conclusions

In the previous pages we have reviewed three main approaches for the detection of chaos in time series: dimensionality, entropy, and the Liapunov exponents. We have reported on the theoretical basis of these concepts and their current fields of applications. We have implemented FORTRAN codes for the numerical measurements and tested them on time series obtained from well–known dynamical systems or stochastic processes. We have then applied the codes to the significant-wave height (SWH) time series. We now summarize the conclusions for each particular approach.

5.1 Dimensional Analysis of Time Series.

5.1.1 Methodological Conclusions.

The distinction betweeen a deterministic and a random dynamics is a pure abstraction: any stochastic process admits a representation as an infinite dimensional dynamical system, with the time forward shift as the law of evolution. A relevant question about a given dynamics is, however, whether it admits an isomorphic (in some specified sense) representation as a finite dimensional dynamical system. The

answer to this question should be a goal for the detection of chaos. Indeed, given the finite length of the data records, we can only hope to be able to distinguish between low–dimensional and high–dimensional dynamics. Of course low–dimensional dynamics is not synonymous with chaos, since there also exists low–dimensional dynamics which are simple. The characterization of such dynamics is often easy (since, for example, the measures supporting periodic and quasiperiodic dynamics are entire–dimensional). Thus, a key feature to distinguish between chaotic and stochastic processes is the dimensionality of the dynamics. The dimension of a measure could be thought of as the *global geometric randomness* of the process.

Under fairly general hypotheses for both dynamical systems and stochastic processes, the orbits (realizations, in the case of stochastic processes) of such systems can be classified according to their long term statistical behaviour, and each class of orbits can be thought of as a probability space endowed with an ergodic probability measure. The dimension of such measures gives an account of the geometric size of the set of orbits displaying a common long–term statistical behaviour. The interest of this type of analysis is quite independent of the deterministic or stochastic nature of the data, since it gives information on a relevant intrinsic feature of the time series in both cases.

The Grassberger–Procaccia algorithm permits us to obtain rather efficient numerical estimates of a particular dimension, namely the correlation dimension ν, which has thus become the form of dimension most widely used in practice. If the series is an orbit of a multivariate dynamical system with an ergodic measure μ, the correlation dimension ν allows a lower bound for the dimension of μ; furthermore, if the dynamical system enjoys suitable mixing conditions, the correlation dimension ν converges to the dimension of μ.

If the time series is a univariate measurement (observable) obtained from a dynamical system (strange attractor hypothesis (SAH), see chapter one), the estimates ν_m of the correlation dimension of the m–dimensional embedding of the series afford lower bounds for the dimension of μ and, in consequence, for the number of variables which can describe the dynamics.

When the series is a realization of an ergodic stochastic process, the estimates

of ν_m provide lower bounds for the dimension of certain projections on $\mathbb{R}^m$ of the invariant measure μ (see section **4.4.1**). Thus it seems plausible that each ν_m is also a lower bound for the number of variables capable of supporting the dynamics.

Since the dimension is an invariant geometric feature for bilipchitz mappings, when we speak of variables supporting the dynamics, we mean a dynamical system isomorphic in a bilipchitz sense to that generating the data, and preserving the measure structure.

Special care is needed to avoid the possibility that the recording noise could increase the estimates of ν_m thus masking the existence of hidden chaotic dynamics. In section **2.2.3** we show how to deal with a white noise of small amplitude. In section **2.3.2** we show how to treat the noise produced by rounded off and truncate data, present in the SWH time series.

To summarize: high values of ν_m rule out the possibility of chaos. Taking into account the fact that convergence to the true values of ν_m could be slow, we may rely for the rejection of the SAH on a steady increase of the estimates of the values of ν_m for increasing m (see **5.1.2** for comments on the practical limitations of the correlation dimension).

5.1.2 Conclusions on the Dimensional Analysis of SWH–Series.

In section **2.3** we analyze two different time series obtained from the raw sea elevation data, the Gi_S and the Gi_{MO} series. Both of them show a steady increase in the estimates of ν_m with m ranging from 1 to 13 (see figures C.28 and C.29). These results are consistent with several surrogate series obtained by adding noises of an amplitude of the order of the recording noise to the original data. On the grounds of section **5.1.1** we can assert that any dynamics generating the SWH series should depend on a minimum number of the order of fifteen variables. In order to test whether this bound may be raised a further recording of data is needed.

5.2 Conclusions on the Analysis of Entropy.

5.2.1 Methodological Conclusions.

The Kolmogorov entropy gives a measurement of the amount of randomness in a given dynamics from a purely measure–theoretic point of view, which is independent of any particular geometric representation of the dynamics. Two dynamics with different entropies cannot be isomorphic under any isomorphism preserving the measure structure of the dynamics. Thus the entropy of a probability measure provides a discriminating parameter which is finer than its dimension. The higher the entropy of a measure, the more random it may be considered. For example, a stochastic process obtained from an i.i.d. sequence of random variables has infinite entropy if the random variables do not have a discrete distribution [Billingsley,78].

The success of the Grassberger–Procaccia algorithm is partly based on the fact that it permits a straightforward computation of the K_2-entropy, a lower bound for the Kolmogorov entropy of a measure. We develop a code for such a computation. In this report we provide a proof showing that the algorithm used in this computation is well founded in the case of dynamical systems (see section **3.2**). It seems plausible that this is also true for ergodic stationary processes (see section **4.4.1**), though we do not have references of the existence of a similar proof in this case.

The use given to the K_2-entropy is very similar to that of the correlation dimension: each scale ε in the analysis of data affords a new estimate $K_2(\varepsilon)$ of the K_2-entropy. A steady increase of such estimates when ε decreases indicates highly random dynamics.

A major drawback of the K_2-entropy, and indeed of the ν_m estimates for the correlation dimension, is that there exist highly random stochastic processes which behave from the viewpoint of the algorithms as low–dimensional/low–random dynamical systems: both the values of ν_m and $K_2(\varepsilon)$ increase so slowly that they seem stabilized at small values (*two* for the correlation dimension, 0.3 for the K_2-entropy). This is the case for stochastic processes with the power spectrum acumulated at low frequencies, and it has been documented in section **3.3.2** for what we have called

slow–cosines processes. We report on the correlation and entropy analysis of some realizations of these kinds of processes in figures **C.48** and **C.50** respectively. Analysis at very small scales is needed to obtain significant increases of the ν_m and $K_2(\varepsilon)$ estimates. The recurrence times required to find neighbouring points at such distances are so large that the computation goes beyond our present capability. The algorithms for computing the K_2-entropy and the correlation dimension are thus useless for analyzing those kinds of data.

5.2.2 Conclusions on the K_2–Entropy Computation for SWH Series.

The results of the computation of K_2–entropies for the SWH series reported in section **3.3.1** indicates a highly random process. The K_2–entropy plots of the SWH process Gi_S can be seen in figure C.38. The degree of randomness, computed up to dimension/time 25, is comparable (although smaller) to that of a Gaussian white noise. The K_2-entropy obtained for the lower scales of resolution in the SWH series is ten times higher than in the Lorenz flow, and three times higher than in the Henon map. Such results are robust with respect to added noises of similar intensity of recording noise. This confirms the conclusions obtained from the dimensional analysis.

A natural goal for the detection of chaos in time series with random behaviour is that of finding some hidden low–dimensional/low–random component in the global signal. A natural way to address this issue is to filter the power spectrum of the process around the frequencies corresponding to the highest peaks of the spectrum. In section **3.3.2** we use low–pass filters with cutoff frequencies slightly higher than the two larger peaks of the spectrum, corresponding to a week and twenty days approximately. We then analyze the filtered series, and bump into the same obstacle we found in the *slow–cosines effect* commented above (see the correlation analysis of the filtered series in figures **C.43** and **C.45**). We thus generate several surrogate series with randomize phases, and find for these series a behaviour similar to that of the original filtered series. On this ground we can only assert that much larger records (or new ideas) would be needed to determine the existence of some hidden

chaotic component in the series.

5.3 Conclusions on Liapunov Exponents.

5.3.1 Methodological Conclusions.

Liapunov exponents for ergodic measures exist both in dynamical systems and in some stochastic processes. In the case of smooth dynamical systems they give information about the rate of divergence of nearby orbits. Whilst the dimension gives a *static measurement* of the *geometric global* degree of randomness of a dynamics, the Liapunov spectrum gives a *dynamic measurement* of the *(geometric) local* degree of randomness of a smooth dynamics (sensitivity to initial conditions). The Liapunov spectrum is related in many cases to the measure–theoretic randomness (entropy) and to the dimension of the measure (see section **B.3.3**), thus closing the circle of ideas reviewed in this report: the entropy is in some sense proportional both to the divergence of nearby orbits (given by the Liapunov spectrum), and to the geometric space occupied by the dynamics (given by the dimension of the measure). Liapunov exponents are diffeomorphic–invariant, and can thus be recuperated from scalar time series generated by smooth dynamical systems. They also provide valuable information for such systems about the local geometric distribution of the invariant measure.

The Liapunov spectrum also exists for some ergodic stochastic processes, namely for random products of matrices. But in their abscence , the above interpretation of the Liapunov exponents as rates of divergence of nearby orbits is meaningless.

We have implemented an algorithm based on sligth variations of the Eckmann–Ruelle algorithm [Eckmann & Ruelle,85] (see section **4.2**): a sequence of matrices is obtained by local linear fits at a given scale ρ to the flow of the orbits. The measurement of the Liapunov spectrum is a more delicate task than that of computing the correlation dimension or the K_2–entropy, which enjoys a nice robustness. We give a detailed account of the significance and use of the relevant parameters which require

a fine–tuned adjustment to obtain correct estimates for the Liapunov spectrum (see section **4.3**). We report on a proof, new to our knowledge, of the fact that such algorithm is convergent when applied to a smooth dynamics, and that the values $\lambda_i(\rho, p)$ of the obtained Liapunov exponents depend both on ρ and on the norm p considered for the fitting process. We give numerical evidence of this dependence in section **4.3.4** (see tables C.1 and C.2).

We propose a new test for chaoticity based upon a measurement of the smoothness of a recorded dynamics. The algorithm we have developed for the computation of the Liapunov spectrum can be applied to any time series. A major difference between the flow generated by a differentiable dynamics and the time–passing flow of a stochastic process is that the former should be well approximated by linear fits, whilst a highly random process could hardly generate a differentiable flow. We introduce a parameter $r(\rho, p)$ that estimates the averaged amount of error in the linear fits (see section **4.4.3**). A small amount of error is a sign of differentiable dynamics. By virtue of the results mentioned above this sign will be reinforced if the Liapunov exponents converge as ρ tends to zero. On the contrary a high amount of error indicates the non-differentiability of the flow. This indication becomes stronger when accompanied by non–convergent Liapunov exponents $\lambda_i(\rho, p)$ for decreasing radii ρ. In section **4.4.3** we provide numerical evidence of these facts, obtained for both differentiable and stochastic dynamics. In particular, when the white noise is uniformly distributed, we find a remarkably precise power scaling law dependence of the Liapunov exponent with respect to the radius ρ of the balls, together with a high amount of error in the linear fits (this can be seen in figure C.54). Furthermore, large values of ρ render a null or even negative value for $\lambda_i(\rho, 2)$, which shows that these exponents do not measure any real rate of orbital divergence. When ρ decreases $\lambda_i(\rho, 2)$ goes to infinity.

5.3.2 Conclusions on the Estimates of the Liapunov spectrum of SWH series.

We compute the Liapunov spectrum of the Gi_S series, for embedding dimensions ranging from 1 to 15. This is reported in tables C.6 to C.13. The values of the

exponents thus obtained are either negative or slightly positive. For dimensions between 1 and 3 the values obtained for the parameter $r(\rho, 2)$ indicate low accuracy. When the radius ρ decreases, the accuracy of the linear fits worsens instead of improving, as it should in the case of a differentiable dynamics. The parameter $r(\rho, 2)$ has been tested for dimensions 1 to 15, and shows low accuracy of the linear fits. On these grounds we can rule out the possibility of existence of a low–dimensional differentiable dynamics. We also rule out the predictability of this process by means of locally linear predictors. The standard methods of locally linear filtering can also be rejected for this procces.

A remarkable feature of the analyzed data is their strong stationarity with respect to the Liapunov spectrum, evidenced both by the independence of the estimated exponents from the starting point of the orbit, and by the number of iterates needed for the convergence of the algorithm (see tables C.7, C.11 and C.8). This shows that the process enjoys nice ergodic properties.

5.3.3 Measurement of the Sensitivity to Initial Conditions of the Gi_S Series.

Liapunov exponents have been estimated with standard algorithms for time series generated by stochastic processes. Such measurements are wrong if they are intended to estimate the rate of divergence of nearby orbits. We propose in section 4.4.4 a new algorithm for solving this problem, based upon a simple modification of the Eckmann–Ruelle algorithm. We test this algorithm for smooth dynamics, obtaining values near the maximum Liapunov exponent, and for various stochastic processes. In this regard we also find a power–law dependence of the uniformly distributed white noise (see figure C.55 and table C.26). We then use this algorithm to estimate the rate of divergence of nearby orbits for the Gi_S series (tables C.22 and C.23). There exists divergence at all scales, with an scaling power law dependence between the radius and the rate of divergence estimated by our algorithm, as in the case of the above mentioned white noise (figure C.56).

5.4 Closing Remark.

Geometric dimension, entropy, and Liapunov exponents are three complementary aspects of an invariant measure. If the underlying generating mechanism of a time series has suitable properties of ergodicity, these three features can be quantified by mean of efficient algorithms. Thus, a valuable information about the analyzed process can be obtained, regardless of its deterministic or random nature. The application of each one of these techniques to the SWH–series has coincided in the conclusion that the generating mechanism is highly random, whilst it has nice ergodic properties. This ergodic behaviour allows us to hope that analysis based in further development of the techniques reviewed in this report will be fruitful. In particular, multifractal analysis, based upon the local dimension (see section **2.2.1**), is focusing a great deal of volume of late research. These features of time series is allowing to obtain a new classification of stochastic processes. It is our belief that this will in turn make possible more ambitious–and difficult–goals, as the nonlinear modelling and prediction.

Appendix A

Further Testing based on the Correlation Integral

In this appendix we include some issues that have not played the main role in the analysis we have developed for SWH–series, though they are relevant to the non–linear analysis of time–series. In particular, we review the estimate of correlation dimension by the method of least–squares. We think this is an important issue, since it has been well–founded that the method works in reasonable contexts [Cutler,91]. An efficient program for computing the correlation dimension is the basis for implementing new tests on chaos detection. We survey some of these tests in a second section.

A.1 The Computation of the Correlation Dimension by the Least–Squares Method.

A standard way used in the literature to get an estimate value $\hat{\nu}$ for the correlation dimension ν (see chapter 2 for definitions) is a *least-squares regression* over the vector range $\mathbf{r} = (r_1, r_2, \ldots, r_k)$ for which the correlation integral was computed. This method has been occasionally criticized [Theiler,90]. It seems, however, that least–squares estimate is the only method which has been proved to converge [Cutler,91].

The commendable paper by Cutler analyzes both the systematic and the statistical errors of the least–square procedure when estimating correlation dimension. [Cutler,91] also analyzes other nearest–neighbor techniques to obtain confidence intervals for estimates of the information dimension.

The analysis of the estimation method stems from the asymptotic statistical theory developed by [Denker & Keller,86] for time series data observed on smooth dynamical systems with 'good' mixing properties. More precisely, in [Denker & Keller,86] it is proved that, if a scanned trajectory in a dynamical system can be represented as a Lipschtiz functional of an absolutely regular (or *weak Bernouilli* in terms of ergodic theory) finite–range valued stochastic process with power–like decaying mixing coefficients, rigorous statistical results on the limit distribution of the Grassberger–Procaccia estimator of the correlation integral

$$C(r) = \int\int 1_{\{(x,y):\|x-y\|<r\}} d\mu(x) d\mu(y)$$

hold. Using this theory, Cutler finds the behaviour of the (least-squares) *systematic error*

$$d(\mathbf{r}) = \hat{\nu}(\mathbf{r}) - \nu, \tag{A.1}$$

which depends on the choice of the vector of scales $\mathbf{r}$ and on the behaviour of the function $\mathrm{wi}(r) = \log C_0(r)$ ($C_0(r)$ as in (2.19)) (see **2.2.2**). The function $\mathrm{wi}(r)$, which is called the *wandering intercept* by Cutler, plays an important role in the analysis. Observe that the hypothesis (2.20)

$$\lim_{r\to 0^+} \frac{\log C_0(r)}{\log r} = 0 \tag{A.2}$$

does not imply convergence of the wandering intercept as r goes to zero.

When $\mathrm{wi}(r)$ converges to some constant ($C(r)$ exhibits 'clean' scaling behaviour) the error vanishes to zero as the vector $\mathbf{r}$ is geometrically compressed by a factor s^k, $0 < s < 1$, that is

$$\lim_{k\to+\infty} d(s^k \mathbf{r}) = 0 \tag{A.3}$$

and the method converges.

If the intercept wi does not converge as r tends to zero, (A.3) is generally false. Observe that the function $\mathrm{wi}(r)$ may oscillate (or even diverge to $\pm\infty$), which can

be considered the most probable case for functions just verifying (A.2). Cutler conjectures that this is unfortunately what generically happens when considering Cantorian geometric structures (as strange attractors typically are). In this situation the choice of **r** (and the way its components tend to zero) becomes crucial, so that when these are adequately chosen the method still can be shown to converge.

On the other hand, the problem of the *statistical error* for the least-square method is by no means obvious. This is because succesive values of $C(r, N)$, say $C(r, N)$ and $C(r+\varepsilon, N)$, are not independent. Therefore the usual error rendered by the least square regression (which assumes uncorrelated observations with common variance) has nothing to do with the true error involved in the dimension estimation. This, in fact, has been the main point against the method of least-squares for estimating ν. The paper by Cutler also solves this problem, using the results in [Denker & Keller,86], by calculating the correct expressions for the involved covariances associated with the correlated observations of orbits in the dynamical system. Independence (the usual requirement for a linear regression model) is no longer necessary in this context as long as the system has good mixing properties (see [Denker & Keller,86] for detailed conditions). Fortunately, Denker and Keller show that many dynamical systems are mixing enough to apply these methods.

A.2 Tests Designed from the Correlation Integral.

The standard statistical tests for 'white randomness' stems from the correlogram and the periodogram analysis (see e.g. [Priestley,81]). From the example of the saw–tooth dynamics (see **1.1.1**), we are aware that a fast–vanishing correlogram and a constant periodogram do not guarantee true randomness. This observation also applies to the case of statistical modelling when a residual analysis (see section **1.1**) is performed to assure the goodness/whiteness of a fitting model for the data. In such cases, a non–pure random residues series could pass the statistical tests as white noise, thus implying that an unfortunate model is accepted as good. Further testing is then highly desirable.

Some non–parametric tests based on the correlation dimension are capable of checking the whiteness of a signal beyond the statistical analysis when applied in the residual analysis. Some of these tests have been able to unmask some unfair models that had passed the standard tests for whiteness. Since a thorough study is beyond the scope of this appendix, we merely point out the ideas behind some of the more popular tests, devised by scientists from economics to astrophysics, which emphasizes the great applicability of these methods.

A.2.1 Phase Randomization Test.

A useful test, originally proposed to discriminate between low–dimensional chaos and colored noise, is *phase randomization* [Osborne et al,86] or the *method of surrogate data* [Theiler,91].

The method consists of creating at random an ensemble (say, ten) of time series with the same power spectrum as the original (therefore the same autocorrelation). A way to do this is to randomize the Fourier phases of the Fourier transform (via FFT) of the primitive series and then to take the inverse Fourier transform. Of course, the output supplied by such trickery is a series having the same power spectrum as the problem series but which has lost the phase information of the latter. Phase information is crucial for the understanding of a chaotic signal from a dynamical system. Therefore if any statistic (dimension, entropy, Liapunov exponents) computed for the surrogate data of the ensemble is significantly different from that computed for the original series, the hypothesis that the original time series is a linearly correlated noise can be rejected. The fact that an ensemble is available provides statistical significance. On the contrary, the invariance of the correlation dimension, or any other statistic, under phase randomization is a powerful suggestion that the series was not generated by low–dimensional chaos. Accepting the hypothesis of deterministic chaos is, as usual, a harder task and cannot be inferred from a drastic change in the log–log plot of the correlation integral for the surrogate data, although this of course helps.

In chapters **2** and **3** we applied the phase–randomization test to several signals,

including SWH–series.

A.2.2 Testing Robustness.

There are a number of tests devised and employed in different applied disciplines in order to obtain reliable estimates of the statistics involved.

A natural approach is to consider a diffeomorphism of the measured signal, or equivalently a diffeomorphism G of the reconstructed set A_m. If the signal was recorded on a strange attractor, estimates of, say, the dimension of the sets A_m and $B_m = G(A_m)$ must coincide according to Taken's theorem. A more robust estimate for the dimension of the original attractor can be obtained if several diffeomorphic sets to the m-dimensional Ruelle–Takens embedding of the data are considered. The estimate can even be improved if a known noise effect of the measure apparatus is damped by adequate smooth mappings G's.

This test has been applied in several circumstances throughout chapter **2** and chapter **3**.

It has been argued [Canizzo et al,90] that the dimension of a stochastic signal with power-law spectra should change drastically if diffeomorphically deformed. This is the basis for a new test for discerning between chaos and colored noise, which enhances the interest of this testing.

We include some further tests related to the correlation integral for the sake of completeness.

A.2.3 Testing Model Building.

Assume that a smooth time series model (e.g. an p–th order autoregressive model, considered the most adequate spectrum estimators) has been fitted to a given data

series, which was picked from an orbit of a chaotic dynamical system. Then, generically, the residual series of the model must have the same (correlation) dimension as the original series.

The above test was called the *Residual Diagnostic for Deterministic Chaos* by its authors [Brock & Sayers,88], who used this test to reject the hypothesis that certain economic series could be fitted by a near unit root autoregressive process. Such processes can generate low dimension estimates, but the Residual Diagnostic eliminates the chance that an autoregressive model simulates the data series correctly if the dimension estimates of the residuals keep at a low level. If, on the contrary, the adjusted model is fair the independent and identically distributed (i.i.d.) residues series will (dimensionally) fill each embedding space, and therefore there will be a jump, in a log-log plot, compared with the estimates for the data.

A.2.4 Shuffling Test.

This is a test for whiteness due to [Scheinkman & Lebaron,89], where it was employed to show strong evidence against the hypothesis of i.i.d. even in the (statistically white) residuals of a non-linear model (ARCH) fitted to financial data. The method consists of the construction of an artificial series by randomly shuffling the original residual series and then comparing the log-log plot of the correlation integral for both series. If the shuffled plot is identical to the plot estimated for the original residues, then this supports the hypothesis of i.i.d. for the residues. The test lies in the fact that a shuffling of the supposedly white residuals destroys any hidden structure and causes a different behaviour between the two correlation plots.

A.2.5 BDS Test.

This is a test designed by the economists Brock, Dechert and Scheinkman (BDS) [Brock et al,87] in order to check the hypothesis of i.i.d. against a wide class of dependence alternatives. BDS created a family of statistics based upon the quantity

$$D^m(r) = C^m(r) - \left[C^1(r)\right]^m, \tag{A.4}$$

where $C^m(r)$ denotes the correlation integral in (2.15). Under the hypothesis of IID $C^m(r) = [C^1(r)]^m$ holds with probability one (as the length of the series tends to infinity) [Brock et al,87]. Therefore i.i.d. is characterized by $D^m(r) \simeq 0$ for all m and all $r > 0$, and the statistics D^m will converge to a non-zero value for some pair (m, r) when it is computed for residuals asymptotically non–i.i.d. (showing the mis–specification of the model associated with the residues).

The interest of the BDS test is at least twofold. First, there exists formal statistical results of a central–limit type [Brock et al,87] for the BDS–statistic D^m as well as results on the (different) behaviour of a BDS–type statistic both for random and deterministic data [Brock & Dechert,91]. Therefore the BDS test allows us to evaluate statistical significance, whereas e.g. the shuffling test does not. Secondly, the method can reveal the existence of a non–linear (deterministic or stochastic) structure where other methods have failed. Note that even when a scaling law for the correlation integral could not exist the BDS test still renders information about the presence of a non–linear structure in the data that could fool the spectrum analysis.

Appendix B

Further Theoretical and Practical Facts on Liapunov Exponents

B.1 Introduction.

If $x_{k+1} = f(x_k)$ is the evolution equation of a dynamical system with a diffeomorphism $f : \mathbb{R}^m \mapsto \mathbb{R}^m$ as the transition function, the behaviour of orbits close to a fixed point x depends on the spectrum of the tangent map $D_x f$. Generically (hyperbolic equilibrium) this spectrum is composed of a set of eigenvalues $\text{Spec}^+ = \{\alpha_1, \alpha_2, \ldots, \alpha_r\}$ with $\mid \alpha_i \mid > 1$, $1 \leq i \leq r$, and a set of eigenvalues $\text{Spec}^- = \{\alpha_{r+1}, \alpha_{r+2}, \ldots, \alpha_s\}$ with $\mid \alpha_i \mid < 1$, $i > r$. The tangent space T_x splits into two subspaces: E_x^+ generated by the eigenvectors with eigenvalues in Spec^+, and E_x^- generated by the eigenvalues in Spec^-. The subspace E_x^+ is the linear subspace tangent to the smooth manifold V_x^u, called *unstable manifold* and defined by

$$V_x^u = \{y \in \mathbb{R}^m : \lim_{k \to -\infty} f^k(y) = x\}.$$

Analogously the subspace E_x^- is the linear subspace tangent to the smooth manifold

$$V_x^s = \{y \in \mathbb{R}^m : \lim_{k \to \infty} f^k(y) = x\},$$

called *stable manifold* (see figure C.57). The spectrum of $D_x f$ determines the dimension of E_x^+ and E_x^-, which in turn coincide respectively with the dimension of

V_x^u and V_x^s. Moreover, in the neighbourhood of x the flow $\{f^k\}_k$ is topologically conjugate of the flow of the linearized system $y_{k+1} = D_x f(y_k)$, so that in such a neighbourhood the orbits can be thought of as homeomorphic deformations of those of the linearized system (see for example [Irwin,80] for details of this theory). This gives a quite accurate picture of the behaviour of the flow about the equilibrium point. Unfortunately this information has only a local character.

The complexity of the global behaviour of chaotic non-linear dynamics frequently stems from the complicated manner in which the stable and unstable manifolds of equilibria fold in a bounded region of the state space (in fact, non-linear chaos is not possible for one or two dimensional continuous-time flows, because the orbits, laying inside curves, cannot fold in a complicated manner without self-intersections). A *tour de force* for the theory of dynamical systems is the understanding of the geometric distribution of the orbits and of the invariant measure beyond the equilibrium points. We shall see how Liapunov exponents are in this setting the most useful tool for this purpose available nowadays, as the natural generalization of the spectrum of the tangent map at equilibria. In section **B.2** we describe the geometric structure of the invariant ergodic measure μ in smooth dynamical systems, showing its close relationship with Liapunov exponents. In section **B.3** we report about the relationships between the Liapunov exponents and the dimension properties of the invariant measure. For a more detailed account on the properties described in sections **B.2** and **B.3** we suggest [Eckmann & Ruelle,85]. Lastly, in section **B.4** we describe and comment on the most generalized algorithms in use for the computation of Liapunov exponents.

B.2 Liapunov Exponents and the Geometry of the Invariant Measure.

B.2.1 Dependence of the Liapunov Exponents on the Invariant Measure.

An annoying feature of Liapunov exponents in smooth dynamical systems (see applications 1 and 3 of the Oseledec theorem in section **4.1**) is their dependence on the measure μ. Although the definitions of Liapunov exponents seems to be independent of any measure, they *exist μ-almost everywhere* for every invariant ergodic measure μ. A unique flow $\{f^k\}_k$ can carry a whole family $\{\mu_i\}_{i\in I}$ of invariant measures as we remarked in chapter **1**. Call M_i the set of points where Liapunov exponents coincide with that of μ_i. Of course, the exponents of different ergodic measures could coincise, but for any two measures μ_i and μ_j such that $M_i \neq M_j$ one has $M_i \cap M_j = \emptyset$, $\mu_i(M_i) = 1$ and $\mu_j(M_j) = 1$, thus showing that μ_i and μ_j are mutually singular measures. In this situation exponents evaluated for starting points at different sets M_i are different and the state space might be partitioned in a family of uncountably many such sets M_i, thus posing the problem of what are in practice the relevant Liapunov exponents. The family of measures given in section **1.4** (see (1.16) for the saw-tooth dynamics provide an example of this situation.

In empirical time series, the Lebesgue measure plays a special role. Any measurement carries certain amount of error. If this error is distributed with some standard probability distribution, which is usually absolutely continuous with respect to the Lebesgue measure, any set with null Lebesgue measure has zero probability of being observable. Therefore the only Liapunov exponents observable in empirical or computational experiments are those corresponding to measures μ_i whose associated set M_i has a positive Lebesgue measure. This ensures that, if we regard a realization of our experiment as a sample in a probability space, there is positive probability of picking a point belonging to M_i as starting point of the scanned orbit. We call such Liapunov exponents *physical Liapunov exponents.*

Let us enumerate some examples of situations which are present in many dy-

namical systems (in the following examples we consider continuous-time flows).

- i) A smooth dynamical system with a *stable equlibrium* x. In this case the Liapunov exponents of x coincide with the logarithms of the absolute values of the eigenvalues of the tangent map at x. Therefore all of them are negative. The ergodic invariant measure is δ_x (unit mass at x), and it can be easily checked that any point in the basin of attraction of x has identical Liapunov exponents as x. Since the basin of attraction is an open set with positive Lebesgue measure, the Liapunov exponents are physically observable. A dynamical system with two stable equilibria x and y is a simple example of a system with at least two sets of physical Liapunov exponents, corresponding respectively to the ergodic invariant measures δ_x and δ_y.

- ii) A *periodic orbit* $\Gamma = \{f^t(x_0) : 0 \leq t \leq T\}$ of period T. If f^t denotes the continuous-time flow, the tangent map $D_x f^T$ at any point x of the orbit is the identity when restricted to the direction tangent to the orbit so that the corresponding eigenvalue is 1, and thus at least one of the Liapunov exponents of each point of the orbit is null. If the orbit is attracting, all of the remaining exponents are negative. The invariant measure is

$$\mu(A) = T^{-1} \int_0^T 1_{A \cap \Gamma}(f^t(x_0))dt, \tag{B.1}$$

  expression which gives the rate of time spent on A by $f^t(x)$ along a whole period T (recall that $1_X(\cdot)$ is the characteristic function of X). This measure is concentrated on the periodic orbit Γ, and the set on which its Liapunov exponents remain fixed is the open basin of attraction of Γ, with positive Lebesgue measure. Reciprocally, Liapunov exponents allow us to characterize attracting periodic orbits: if all of the Liapunov exponents of a measure μ are negative except for one which is null, then the measure is that described in (B.1) (see [Eckmann & Ruelle,85]).

  The following example enlightens some aspects of the dynamics of strange attractors.

- iii) For orbits of *continuous-time dynamical systems*, $\dot{x} = F(x)$, in a compact manifold M there exist two possible situations:

  a) The orbit starts in the basin of a fixed point x, as analyzed in i).

b) The orbit does not terminates in a fixed point but wanders perpetually around M. This is the case of the orbits of a chaotic attractor. In this case there always is at least one Liapunov exponent equal to zero. This can be seen by means of the following heuristic argument, which extends the argument given above for periodic orbits to the case of recurrent (non-wandering) orbits. Since the measure μ is ergodic, μ-almost any x generates a recurrent orbit: we can find arbitrarily large periods of time T for which $f^T(x)$ is arbitrarily close to x. The tangent map $D_x f^T$ sends the vector $F(x)$ to the vector $F(f^T(x))$ (observe that they are tangent vectors to the curve $f^t(x)$ at $t = 0$ anf $t = T$ respectively), and the flow f^T maps the curve f^t onto the curve f^{t+T}. Since $f^T(x)$ can be taken arbitrarily close to x and $F(x) \neq 0$, the tangent map restricted to the tangent direction to the orbit is arbitrarily close to the identity after arbitrarily long periods of time, thus showing that the Liapunov exponents along the direction $F(x)$ is zero (for a complete proof see [Haken,83]). This implies that orbits of chaotic attractors always have one null Liapunov exponent, as can be seen in the case of Rossler and Lorenz attractors in the numerical estimates obtained in **4.3**.

If the adherence of an orbit of a continuous-time smooth flow does not reduce to a fixed point the above reasoning remains valid, and there exists a null exponent. In consequence, if none of the Liapunov exponents along an orbit is zero, the adherence of the orbit reduces to fixed points so that such an orbit is either that of a fixed point (which will be stable in the case where all of the Liapunov exponents are negative) or an orbit starting in its basin of attraction. For discrete-time flows there does not exist such a result. For example, the Henon map (see section **4.3**) has two Liapunov exponents, one positive and one negative.

B.2.2 Stable and Unstable Manifolds of Orbits in Chaotic Attractors.

We shall see how the Oseledec theorem permits the extension of the concepts of stable and unstable manifolds (see section **B.1**) beyond the equilibrium points. This theory can be developed both for discrete and continuous time smooth flows in a

smooth manifold M. We assume that an ergodic measure μ has been fixed, and the Liapunov exponents are those associated with this measure. Then for any perturbation vector v of the subspace $E_i(x)$ spanned by the eigenvectors corresponding to the Liapuñov exponents smaller than or equal to λ_i we have

$$\lim_{n\to\infty} n^{-1} \log \| D_x f^n(v) \| \leq \lambda_i(x).$$

Assume that λ_i is negative. What happens if we look at the real divergence between orbits, instead of the image of the vector v under $D_x f^n$?. If we fix a $\lambda > \lambda_i$ it can be proved that for a sufficiently small $\varepsilon > 0$ the set

$$V_x^s(x,\varepsilon) = \{y : d(f^t(x), f^t(y)) < \varepsilon \exp(\lambda t) \text{ for all } t \geq 0\}$$

is, for μ-every x, a piece of a smooth manifold contained in a ball centered at x and with radius ε. Furthermore it is tangent to and with the same dimension as $E_i(x)$, provided $\lambda_{i-1} > \lambda$. In the case of a flow defined for negative values of t, the following definition of *global stable manifold* makes sense

$$V_x^{i,s} = \bigcup_{t>0} f^{-t}(V_x^s(\lambda,\varepsilon)).$$

It is obtained by gluing together the pre-images of a local stable manifold. This can be shown to be equivalent to

$$V_x^{i,s} = \{y : \lim_{t\to\infty} \frac{1}{t} \log d(f^t(x), f^t(y)) \leq \lambda_i\},$$

(compare with the definition given in section **B.1** for equilibria). When λ_i is taken as the largest negative Liapunov exponent (i.e. $\lambda_{i-1} \geq 0$) the global stable manifold V_x^s is called the *stable manifold*. For invertible flows, one defines local unstable manifolds $V_x^u(\lambda,\varepsilon)$, global unstable manifolds $V_x^{i,u}$ and the *unstable manifold* V_x^u just by reversing the signs of the time in the above definitions of stable manifolds. See [Ruelle,79] for a complete proof.

Let us discuss the issue of the distribution of the ergodic measure μ on the manifold M supporting the flow. Global manifolds may be smooth curves, surfaces, etc. As we have remarked in section **B.1**, when the manifold M supporting the flow is compact, global manifolds are constricted to fold upon themselves in complicated patterns, since they are contained in bounded regions. The invariant measure should be expected to behave in different ways in stable and unstable manifolds, for whilst

the unstable manifold, tangent to the subspaces spanned by eigenvectors with positive Liapunov exponents, is submitted to an action of *stretching* by the flow, the stable manifold is, on the contrary, submitted to a *shrinking*. We should expect the ergodic measure μ restricted to unstable manifolds, to be absolutely continuous with respect to the Lebesgue measure, and to be singular (concentrated on sets of null-Lebesgue measure) when restricted to stable manifolds. The following example supports this conjecture: if M is the unit circle, the expanding map $f(x) = 2x$ admits the Lebesgue measure as an invariant measure, with $\log 2 > 0$ as the unique Liapunov exponent. The unstable manifold coincides with the whole circle M. On the contrary, the contracting map $f(x) = .5x$ has the measure δ_0, singular with respect to the Lebesgue measure, as the unique ergodic measure with Liapunov exponent $-\log 2 < 0$, and the whole circle coincides in this case with the stable manifold.

The above reasoning gives rise to the concept of SRB measures (SRB for Sinai, Ruelle and Bowen), which are those ergodic measures absolutely continuous with respect to the Lebesgue measure on unstable manifolds. Notice, from the definition of the unstable manifold, that $y \in V_x^u$ if and only if $x \in V_y^u$. Thus, unstable manifolds can be thought of as equivalence classes partitioning the manifold M (unstable foliation). Assume now that we analyze the orbits of a chaotic attractor A with fractional dimension. Then we should expect that the measure on the stable manifolds of points of A (transversal stable foliation) behave as a fractal measure.

Unfortunately, given a smooth dynamical system it is not known whether it carries an ergodic SRB measure. This holds at least for the important family of Axiom-A dynamical systems (see [Arnold,80] for definitions and properties of Axiom-A systems), where the stable and unstable foliations can be well understood. For attractors occurring in Axiom-A systems it can be shown that the SRB is a *physical measure*, in the following sense: if μ is an ergodic measure the Birkhoff ergodic theorem guarantees that for μ-a.e. x, the measure of a μ-measurable set A is given by

$$\mu(A) = \lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} 1_A(f^k(x)) = \lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} \delta_{f^k(x)}.$$

This means that the measure μ coincides with the asymptotic relative frequency of the visits to the set A of the orbit starting at x. If the measure μ is singular with respect to the Lebesgue measure, it could be unobservable from a physical point

of view, since physical experiments only pick starting points x if they belong to a set with positive Lebesgue measure (this issue has been discussed yet for Liapunov exponents). Only in the case that the above formula holds for every x in some set of positive Lébesgue measure can we expect the measure μ to be observable, and this is just what happens for the SRB measures in Axiom-A systems.

B.2.3 Consequences on Numerical Estimates of Liapunov Exponents.

Systems like Lorenz's, which plays a relevant role in the literature of chaotic attractors and provides a classical testing bench for algorithms for numerical computations of Liapunov exponents, also share the property of Axiom-A systems [Eckmann & Ruelle,85]. This has the following consequence: any algorithm for computing the Liapunov exponents of a physical measure based upon a finite data set find the data spatially distributed according to the weight of each region of the attractor with respect to the physical measure. Thus the data set distributed on stable manifolds should be rarified with respect to those on unstable manifolds. But the scanning of nearby orbits with starting points at the stable manifold gives the negative exponents, whilst the orbits with starting points at unstable manifolds give positive Liapunov exponents. Thus a lack of data should be expected to cause bad estimates in negative Liapunov exponents, whilst positive Liapunov exponents should enjoy a strong robustness for numerical estimates. See in section **4.3.3** for numerical evidence of these facts.

B.3 Liapunov Exponents, Dimension and Entropy.

The theory about relationships between Liapunov exponents and different dimensions associated with dynamical systems is rather extensive. Since a detailed description of the known results exceeds the scope of this work we will select and summarize some significative aspects.

B.3.1 Some Basic Parameters.

Let $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \ldots$ be the Liapunov spectrum of an ergodic measure μ in a manifold M. We take in the sequence every exponent repeated according to its multiplicity. Consider the linear interpolation of the function that assigns every natural number $j \leq \dim M$ the sum of all Liapunov exponents λ_i with $i \leq j$. This is a concave real function $\Lambda(\xi)$ with a unique root $\xi = \mathrm{Dim}_\Lambda \mu$ (see figure fig C.58) which is called *Liapunov dimension* of μ. Another significative parameter is L^+, the value where the maximum of $\Lambda(\xi)$ is attained. It is readily seen that L^+ coincides with the cardinality of all positive Liapunov exponents in the above sequence.

B.3.2 Bounds for Dimension from Liapunov Exponents.

We now remark a notable property of the unstable manifolds. Call *attracting* a set A if it is contained in some open set U contained in turn in the basin of attraction of A, in such a way that orbits starting at U converge to A. More precisely, for $t > 0$ $f^t(U) \subseteq U$ and $A = \bigcap_{t>0} f^t(U)$. Then the unstable manifold of the points of an attracting set contained in A is also contained in A. This is easily seen, since when the time goes backwards, every orbit of a point x in the unstable manifold became arbitrarily close to some point of the attractor, and this shows that it gets inside U in a finite time T. Therefore for $t > T$ there is some $z \in U$ with $x = f^t(z)$, and this shows that x belongs to $\bigcap_{t>T} f^t(U) = A$. Notice that, although some definitions of chaotic attractors avoid the requirement of existence of an open neighbourhood of the attractor contained in its basin of attraction, this situation is likely to hold, in attractors observable from real data.

The above reasoning shows that the Hausdorff dimension of A is bounded below by that of the unstable manifold. If μ is an SRB measure then the dimension of the unstable manifolds V_x^u, $x \in A$, affords a lower bound for $\dim\mu$, since μ restricted to V_x^u is absolutely continuous with respect to the Lebesgue measure on the manifold, whose dimension is just that of V_x^u. Now, as we have remarked in **B.2.2** the dimension of V_x^u is that of the subspace spanned by all the eigenvectors with positive Liapunov exponents, i.e. L^+. This shows that $L^+ \leq \dim\mu$.

The function $\Lambda(i)$ gives, for entire i, the exponential rate of growth of the volume elements of dimension i (see **4.1**). Since $Dim_\Lambda \mu$ is a root of $\Lambda(\xi)$, the elements of volume with entire Hausdorff dimension d greater than $\mathrm{Dim}_\Lambda \mu$ are contracted. This allows to expect that $\mathrm{Dim}_\Lambda \mu \leq d$, a conjecture that has been proved by Ilyashenko [Ilyachenko,83]. It has been conjectured [Frederickson et al,83] that for SRB measures $\mathrm{Dim}_\Lambda \mu = \dim \mu$. This relationship is true for C^2-diffeomorphisms on two dimensional manifolds, but there exist also counterexamples (which of course are not SRB measures). See [Grassberger & Procaccia,83c] for a discussion on the type of dynamic mechanism which violates the Kaplan–Yorke conjecture.

B.3.3 Liapunov Exponents and Entropy.

Lastly we summarize some relationships between Liapunov exponents of an ergodic measure and the amount of randomness associated with the measure, quantified by its entropy (see chapter **3** for the construction of measure–theoretic entropy).

From the above definitions we know that $\Lambda(L^+)$ gives the sum of all positive exponents. The following result of Ledrappier and Young [Ledrappier & Young, 84] gives valuable information about the structure of an ergodic measure with compact support for a twice differentiable diffeomorphism on an m-dimensional manifold: such an ergodic measure is SRB if and only if $h(\mu) = \Lambda(L^+)$. Under the specified hypothesis, the computation of all positive Liapunov exponents permits by this formula an estimation of the uncertainty (and in consequence of the forecastability) in the system, as measured by the entropy.

There exist several formulae relating Liapunov exponents, entropy $h(\mu)$ and $\dim \mu$. For example, for the saw-tooth dynamics in chapter **1** (see **1.1.1**) Young [Young,82]) proved that

$$\dim \mu = h(\mu)\lambda^{-1},$$

where μ is any of the Bernouilli measures defined in (1.16), and $\lambda = \log 2$ is the Liapunov exponent of μ. Young [Young,82] showed that $\dim \mu = h(\mu)(\lambda_1^{-1} + |\lambda_2|^{-1})$ holds for C^2-diffeomorphisms on two-dimensional manifolds with an ergodic measure with compact support and Liapunov exponents $\lambda_1 > 0$ and $\lambda_2 < 0$. The general

structure of such relationships suggests the following heuristic principle. The randomness of the measure μ (as measured by the measure-theoretic entropy $h(\mu)$) grows with the rate of divergence of the orbits, and also with the dimension of the space that the orbits have available to spread out. A narrow geometric space eases the finding of the location of future orbits. Estimation of randomness or forecastability based exclusively on Liapunov exponents, regardless of the geometry and dimension of the measure could be meaningless, especially in the case of stochastic processes, where the dimension of the measure could be infinite.

B.4 Other Algorithms for Computing Liapunov Exponents and their Differences with the Algorithm we propose in this Report.

There exist two main families of algorithms for computing Liapunov exponents. The first one attempts to estimate the largest Liapunov exponent, and operates by tracing orbits of nearby points. This family is represented by the algorithm proposed by Wolf, Swift, Swinney and Vastano [Wolf et al,84] and by that proposed by Sato, Sano and Sawada [Sato et al,87]. We describe these algorithms in section **B.4.1**. The second family is represented by the algorithms proposed independently at about the same time by Eckmann and Ruelle [Eckmann & Ruelle,85, Eckmann et al,86] and by Sano and Sawada [Sano & Sawada,85], and it is based upon linear fits of the maps tangent to the flow. Since our algorithm is based upon slight modifications of the algorithm of Eckmann and Ruelle, which has been described in detail in sections **4.2** and **4.3**, we shall devote section **B.4.2** to commenting on algorithm of Sano and Sawada as well as the modifications proposed by Abarbanel, Brown and Kennel [Abarbanel et al,91]. In section **B.4.2** we comment on the *differences* between our algorithm and the algorithm of Eckmann and Ruelle.

B.4.1 Algorithms for Computing the Largest Liapunov Exponent.

a) The algorithm of Wolf, Swift, Swinney and Vastano [Wolf et al.,84].

This algorithm, which we shall abbreviate to WA, is designed to compute the largest Liapunov exponent. It is based on the following idea: given an orbit $\{x_i\}_{1\leq i\leq n}$ in a strange attractor of a smooth flow, tangent vectors tend to align, under the action of the flow, with the direction corresponding to the largest Liapunov exponent, which we shall call *most unstable direction.* The orbits of two nearby points x_0 and y_0 are traced during a period of time t_0 not so large that the length vector $v_1 = f^{t_0}(x_0) - f^{t_0}(y_0)$ is of the order of the total size of the attractor (in order to avoid the folding action of the flow), but sufficiently large as for the direction of v_1 to be approximately tangent to the most unstable direction at the point $x_1 = f^{t_0}(x_0)$. Then we perform a *replacement* of the point y_0 with a data point y_1 which is lying near x_1, and such that the vectors $y_1 - x_1$ and v_1 have approximately the same direction. Then

$$t_1^{-1} \log \frac{d(f^{t_1}(y_1), f^{t_1}(x_1))}{d(y_1, x_1)}$$

gives an estimate of the exponential rate (averaged over the period t_1) at which the flow stretches the vectors of the tangent space at x_1 along the most unstable direction. We then perform N consecutive replacements as above, with N sufficiently large so that the whole attractor is visited by the points $f^k(x)$, $1 \leq k \leq \sum_{i=1}^{N} t_i$. Then

$$\left(\sum_{i=1}^{N} t_i\right)^{-1} \sum_{k=1}^{N} \log \frac{d(f^{t_k}(y_k), f^{t_k}(x_k))}{d(y_k, x_k)}$$

gives an estimate of the largest Liapunov exponent. The subspace spanned by a linearly independent system of k vectors tends, under the action of the flow, to the subspace spanned by the directions corresponding to the k largest Liapunov exponents. A process of replacements as above would give us the rate of growth of k-elements of volume, which is known to coincide (see **4.1**) with the sum of the k largest Liapunov exponents, and from here the respective Liapunov exponents can be computed.

The WA algorithm has been, probably, the most extensively applied to empiri-

cally recorded time series, where the main goal is often to detect chaotic dynamics, giving some estimate of the sensitive dependence on initial conditions for the system, for which the largest Liapunov exponent is needed. The authors [Wolf et al,84] claim that the algorithm works in many cases reasonably well for the computation of the two largest exponents. In fact, [Brock,86] proves that the algorithm converges in the case of smooth dynamics when the lenght of the orbit tends to infinity.

Vastano and Kostelich [Vastano & Kostelich,85] show that the largest Liapunov exponent estimated from the WA algorithm is, in some cases, stable when an attractor is embedded in higher dimensional spaces whilst the algorithm of Eckmann et al. (EA for the sequel) [Eckmann et al,86] (see **4.2** and **B.4.2** below) is not. In the example they exhibit, when applying the EA algorithm, the largest Liapunov exponent is what has been called a spurious Liapunov exponent in the subsequent literature. This spurious exponent is, as it is in many cases (see [Eckmann et al,86] and [Abarbanel et al,91] for some examples), an entire multiple (approximately twice in this case) of the true largest exponent which is also computed by the EA algorithm. In **4.3.5** the causes which motivate the existence of spurious exponents are explained, and methods for their elimination are pointed out.

The WA algorithm could be free of spurious exponents in reconstructions of attractors because it evaluates the stretching of the flow for nearly tangent vectors obtained from data points, which are lying approximately in the images $J(T_y)$ of the tangent spaces under the embedding diffeomorphism J, where the true Liapunov exponents are known to work, whilst the EA algorithm computes the Liapunov exponents from the evolution of arbitrary sets of vectors of the tangent space of the reconstructed space. On the other hand this presents the advantage of getting free of the process of replacement. This allows in many cases to obtain good estimates of the whole Liapunov spectrum, so giving a more complete information than the WA algorithm about the invariant measure and associated dynamics.

b) The algorithm of Sato, Sano and Sawada [Sato et al.,86].

The above authors propose a method to compute the largest Liapunov exponent avoiding the noisy process of replacement.. They fix a data point x, find the nearest

neighbour y and choose a suitable t such that $v_x = f^t(y) - f^t(x)$ is approximately tangent to the most unstable direction at $f^t(x)$. They also fix a short period of time τ such that the action of the τ-time flow $f^\tau(x)$ is approximately linear over vectors in the range of size of v_x. They then estimate the stretching $\lambda(x,t,\tau)$ of the flow in the most unstable direction at x by $\tau^{-1}\log(\| f^\tau(v_x)\| / \| v_x \|)$. Since the most unstable direction at each point is invariant under the flow, they proceed to average $\lambda(x,t,\tau)$ with respect to the invariant measure over the whole attractor. They show numerical evidence that

$$\lambda(t,\tau) = \int \lambda(x,t,\tau) d\mu(x)$$

gives good estimates of the largest Liapunov exponent, at least in some of the best known chaotic attractors, even when they are embedded in higher dimensional spaces. The success of this method depends upon the search of the nearest neighbour of many points, a process wich is rather time consuming. This allows to obtain a good aligment of v_x with the most unstable direction for values of t sufficiently small as to $\lambda(x,t,\tau)$ to give a reasonably accurate estimate of the stretching of the tangent map $D_x f$ in that direction. Since the data points tend to be lying on the unstable manifold of the flow, it is likely that the nearest neighbours are lying near the most unstable directions in many cases, which could explain the remarkable efficiency of the method when working for low values of the parameter t in the computation of the exponents of a Lorenz attractor (this also operates indeed for the WA algorithm).

B.4.2 Algorithms Based upon Linear Fittings.

a) The algorithm of Eckmann and Ruelle (EA) and its modifications [Eckmann & Ruelle,85, Eckmann et al.,86].

Since our algorithm is based upon slight modifications of this algorithm, we do not need to describe it here; this has been done in detail in section **4.2**. We describe here other versions of this algorithm.

Sano and Sawada [Sano & Sawada,85] proposed also about the same time as Eckmann and Ruelle an algorithm based upon linear fits obtained from the data, but they do not use the Q-R algorithm (see **4.2**) to draw from the obtained matrices

the Liapunov exponents. If $S_1, S_2, \ldots, S_N$ is the sequence of matrices, they construct a sequence $B_1, B_2, \ldots, B_N$ of orthonormal basis, in such a way that the vectors of B_k are obtained performing the Gramm–Schimdt process of orthonormalization over the system of vectors $S_k(B_{k-1})$. Thus, the first vector e_k^1 of B_k is the normalized image of the first vector of B_{k-1} under S_{k-1}, and when k increases these vectors tend to align with the most unstable direction. Thus we can draw our approximation to the largest Liapunov exponent from the expression

$$(N\tau)^{-1} \sum_{k=1}^{N} \log \| S_k(e_k^1) \|,$$

where τ is the time between two consecutive points of the orbit. When performing the Gramm–Schmidt orthonormalization the two first vectors of the orthonormalized basis are always contained in the subspace generated by the two first vectors of the original basis. Thus, the subspace spanned by the two first vectors e_k^1 and e_k^2 of B_k tends to be tangent to the subspace spanned by the two most unstable directions of the tangent space at the k-th point of the orbit. Moreover, since e_k^2 is orthogonal to the most unstable direction, it must be aligned with the eigenvector corresponding to the second largest Liapunov exponent (these eigenvectors are always orthogonal since, since the matrices Λ_x are symmetric, see 4.1 for definitions). The reasoning can be generalized to the i-th vector e_k^i, and

$$(N\tau)^{-1} \sum_{k=1}^{N} \log \| S_k(e_k^i) \|, \quad 1 \leq i \leq N,$$

give estimates for the whole Liapunov spectrum. This method is, theoretically, as efficient as the EA algorithm is. In fact the process of orthonormalization proposed by Sano and Sawada is essentially equivalent to the EA method. Moreover, it is easily seen that, if QR is the Q-R decomposition of a matrix A then the matrix Q is just the orthonormal basis obtained from the columns of A by the Gramm–Schmidt process, which shows that the EA algorithm and the algorithm of Sano and Sawada only differ in the routines used and in the way in which the Liapunov exponents are finally read.

We now comment on some modifications proposed by Abarbanel, Brown and Kennel [Abarbanel et al,91]. First they perform the linear fits taking into account terms higher than the first in the Taylor expansion of the flow. A third order fit,

leading to a linear system, seems to improve significantly the accuracy of the fits, although the number of linear equations to be solved increases exponentially with the order of the fit. Second they use the symmetrized matrices $(S^{(n)*} \circ S^{(n)})^{1/2}$ (see **4.1**) appearing in the Oseledec theorem rather than the matrices $S^{(n)}$, as in EA algorithm. In section **4.1** we sketch a proof that this is not necessary.

Lastly, they propose the following modification in the use of the Q-R algorithm, which they use to compute directly the eigenvalues of the symmetric real matrix $A = (S^{(n)*} \circ S^{(n)})^{1/2}$: the Q-R decomposition of A is performed (factorwise, as in the EA algorithm), obtaining $A = Q_1 R_1$. The Q-R decomposition of A is then performed, after having written A in the basis given by the columns of Q_1, i.e. we compute the Q-R decomposition of the matrix $A_2 = Q_1^{-1} A Q_1$ obtaining $A_2 = Q_2 R_2$, and the process continues this way, with the recurrent formulae $A_k = Q_{k-1}^{-1} A_{k-1} Q_{k-1}$ and $A_k = Q_k R_k$. From these identities it follows that $A_k Q_{k-1}^{-1} = Q_{k-1}^{-1} A_{k-1}$, which recurrently applied yields

$$A_k Q_{k-1}^{-1} Q_{k-2}^{-1} \dots Q_1^{-1} = Q_{k-1}^{-1} Q_{k-2}^{-1} \dots Q_1^{-1} A.$$

The formula

$$Q_k^{-1} Q_{k-1}^{-1} Q_{k-2}^{-1} \dots Q_1^{-1} A^k = R_k R_{k-1} R_{k-2} \dots R_1$$

is true for $k = 1$. If we suppose it is true for $k-1$ then

$$R_k R_{k-1} R_{k-2} \dots R_1 = Q_k^{-1} A_k Q_{k-1}^{-1} Q_{k-2}^{-1} \dots Q_1^{-1} A^{k-1} =$$

$$= Q_k^{-1} Q_{k-1}^{-1} Q_{k-2}^{-1} \dots Q_1^{-1} A^{k-1}.$$

This shows that, if $R^{(k)} = R_k R_{k-1} R_{k-2} \dots R_1$ and $Q^{(k)} = Q_1 Q_2 \dots Q_k$, then $Q^{(k)} R^{(k)}$ is the Q-R decomposition of A^k. Now, if A is a symmetric real matrix, and the Q-R algorithm is performed with powers $A \circ A \circ \dots \circ A$, a similar reasoning to that used above to explain the algorithm of Sano and Sawada, shows that the orthonormal matrices Q such that $Q^{(k)} R^{(k)} = A^k$, with $R^{(k)}$ upper triangular, converge to a matrix whose columns are the eigenvalues of A. This requires that Q_k converges to the identity. Moreover in the above algorithm the matrix A_k is the matrix A written in the basis given by the columns of $Q^{(k-1)}$. If these are the eigenvectors of A, A_k must take diagonal form asymptotically, and then the eigenvalues of A_k, which are the same as those of A, can be read in the principal diagonal of A_k.

The method described carries the handicap of an overloading of the available memory, since all the matrices obtained by the linear fits must be recorded to perform the Q-R decomposition of A, and this computation has to be repeated several times. The rate of convergence of this process depends upon the difference between the consecutive eigenvalues of A. If two of them were nearly the same, the process could converge rather slowly, and many iterations could be required. The EA algorithm obtains with only one iteration the same eigenvalues. Moreover, the Q-R algorithm can be performed directly with each linear fit when they are computed, and only the diagonal elements of the 'R' matrices have to be recorded, thus saving a considerable amount of memory.

The use of the Q-R algorithm proposed by Abarbanel, Brown and Kennel gives as a by-product the eigenvalues of the matrix $(S_x^{(n)^*} \circ S_x^{(n)})^{1/2n}$. These eigenvalues converge to the Liapunov exponents for large n but, as the above–mentioned authors have pointed out, this does not hold for small n, even if they are averaged for orbits with starting points taken over the whole attractor. They interpret these averaged eigenvalues as estimates of the rates of divergence of orbits over short periods of time (a sort of predictability behaviour of the system in the short term), and they call them *local Liapunov exponents* (see [Abarbanel et al,91]). A perturbation v of the orbit with starting point x grows at a mean exponential rate over a period of time τn given by $(\tau n)^{-1} \log \| S_x^{(n)}(v) \|$, which we know to be equivalent to

$$(\tau n)^{-1} \log \| (S_x^{(n)^*} \circ S_x^{(n)})^{1/2}(v) \|$$

(see **4.1**). Thus a consistent definition of local Liapunov exponents should be based upon estimates of these quantities. Unfortunately, the spectrum of the matrix $(S_x^{(n)^*} \circ S_x^{(n)})^{1/2}$, which is what the above algorithm actually computes, only gives such estimates for large n.

b) Differences between the EA and our algorithm.

We now comment on the differences between the algorithm of Eckmann and Ruelle and ours. The main difference is that we do not limit the maximum number of data points to perform the linear fits of the flow, but we use instead all the available data points inside balls with a given radius ρ centred at the points of the orbit. As we

remarked in section **4.3**, if there are k of these points, the number of arithmetical operations needed for a least square fit is the order of k^2, that is, the same order as the CPU-time we save by avoiding the ordering of the neighbours by increasing distances to the centre of the ball, operation which is usually performed when limiting the number of points in the linear fits with view to the selection of the nearest neighbours (the main advantage of fits based upon a restricted number of data points is that we can select closest neighbours). In this way we obtain a reasonably efficient algorithm. The main advantage of our proposal is that it lets the invariant measure distribute freely in the balls, which gives more accuracy in the estimates of the negative Liapunov exponents and improves also the bad determination in the matrices for the linear fits. In the case of a lack of data our method could, however, require too large values of ρ to guarantee the minimum of data points needed to obtain linear fits in each ball, thus enhancing the non-linear terms in the Taylor expansion of the flow. The results mentioned in **4.3.3** give a theoretical support to our method, showing that our algorithm converges to the true Liapunov exponents in the sense there specified, and that, in any case, there exists Liapunov exponents for linear fittings in balls of fixed radius ρ. We show that there exist a dependence of these exponents upon the p-norm used in the fits, and how this norm could be used to improve the always problematic estimates of the negative Liapunov exponents.

Other differences between the EA algorithm and ours are that we work in a vectorized form, which allows the direct analysis of data numerically generated by n-dimensional dynamical systems and time series. On the other hand, we use the two–dimensional box–assisted technique for the search of neighbours described in section **2.2.4**.

Appendix C

Figures and Tables

In this last appendix we include all the figures and graphics referenced throughout the report. We developed FORTRAN codes for all the computations specified in the text, that were executed in a CONVEX computer. We ocassionally used some routines from the I.M.S.L. for some standard calculations (these have been referred at the right moment in the text). All the plots were performed by GNUPLOT software available in a CONVEX machine, and printed on HP laser printer. We include the computation time (CPU time) for some of the relevant calculations.

C.1 Testing Laboratory of Chaotic Dynamical Systems.

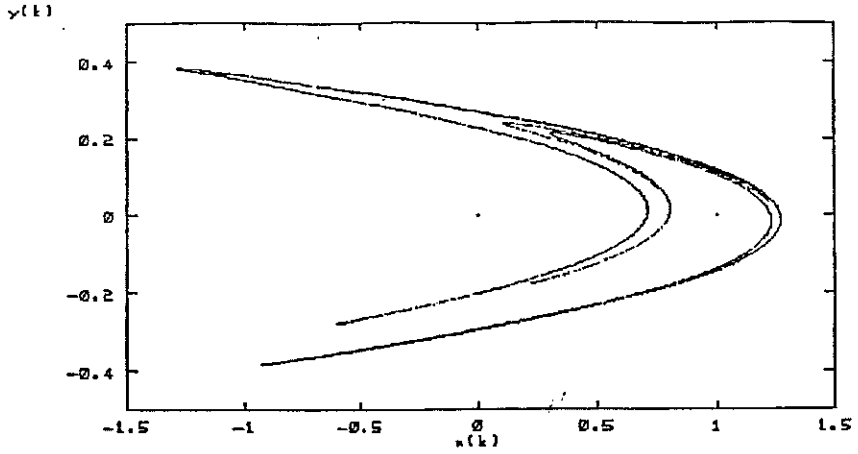


Figure C.1: *Henon system.* Trajectory of the 2-dimensional discrete dynamical system steered by the Henon mapping $f(x,y) = (1 - ax^2 + y, bx)$ [Henon,76], for the standard values of the parameters: $a = 1.4$, $b = 0.3$.

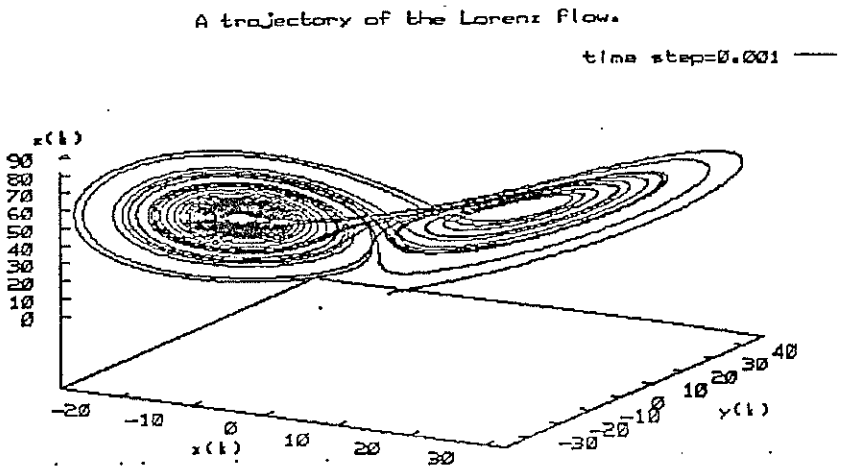


Figure C.2: *Lorenz flow.* Trayectory from the Lorenz system [Lorenz,66]

$$\begin{cases} \dot{x} = -\sigma(x+y) \\ \dot{y} = -xz + rx - y \\ \dot{z} = xy - bz \end{cases}$$

with the standard parameters values of $r = 45.92$, $b = 4$, $\sigma = 16$. Runge–Kutta integration available from the I.M.S.L. was used to generate the displayed orbit, with an integration step of 10^{-2}.

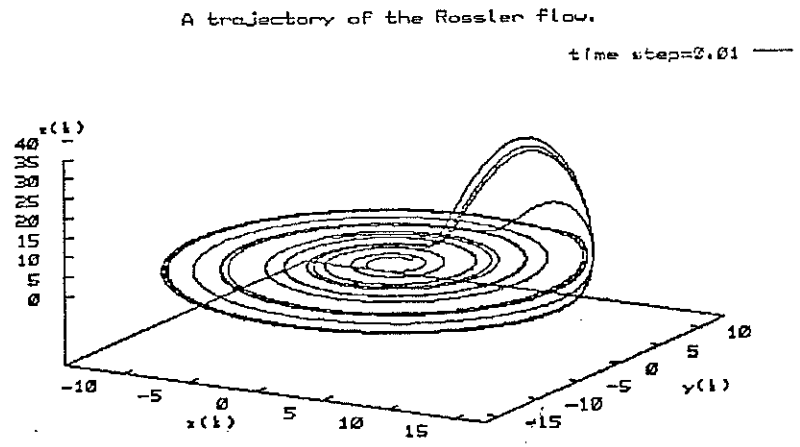


Figure C.3: *Rössler flow.* Plot of an orbit generated by a Runge–Kutta scheme (integration step 10^{-2}) for the Rössler system [Rössler,76],

$$\begin{cases} \dot{x} = -z - y \\ \dot{y} = x + ay \\ \dot{z} = b + z(x - c) \end{cases}$$

for the values of the parameters $a = 0.15$, $b = 0.2$, $c = 10$.

C.2 Power Spectra of Some Deterministic and Noisy Processes.

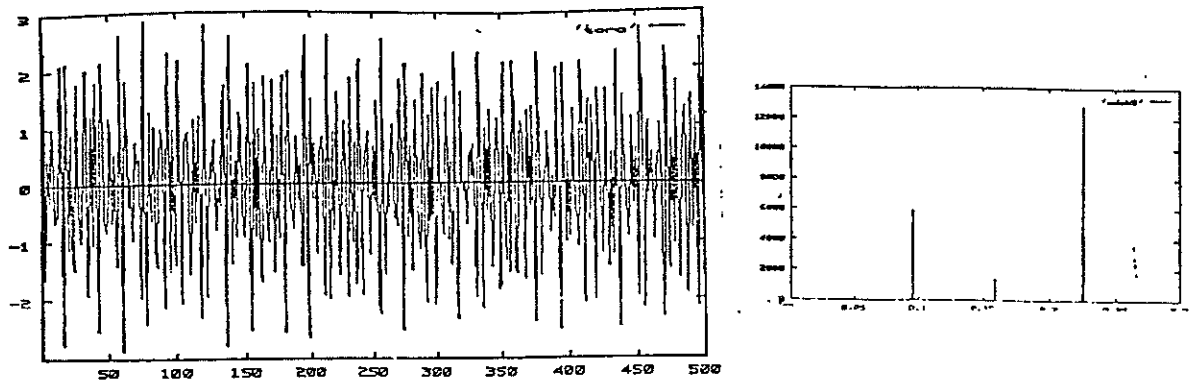


Figure C.4: *Quasiperiodic evolution.* Time evolution and crude power spectrum (computed via FFT-routine of the I.M.S.L.) of a quasi–periodic time evolution, generated by three superimposed frequencies. This kind of temporal evolution can be unmasked from the spectral analysis.

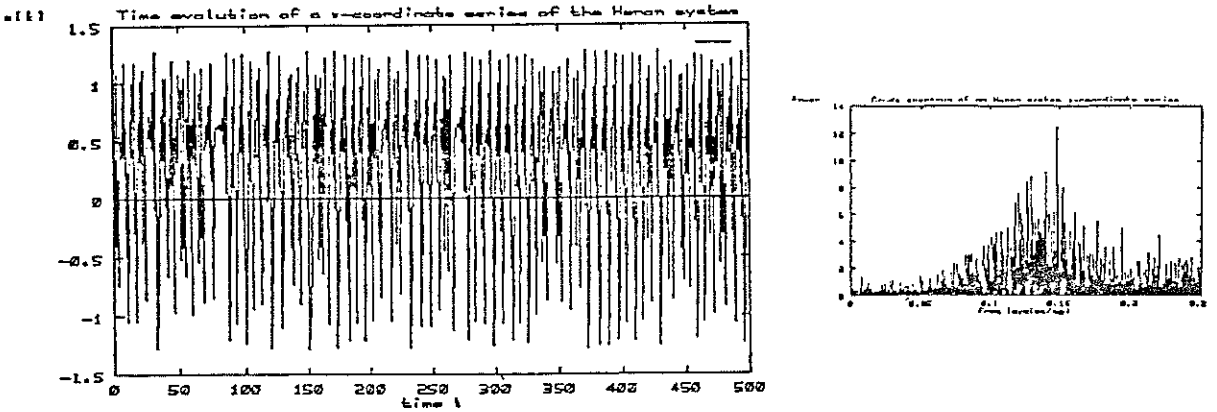


Figure C.5: *Chaotic evolution.* Crude power spectrum and time evolution of a x–coordinate series recorded from the Henon system.

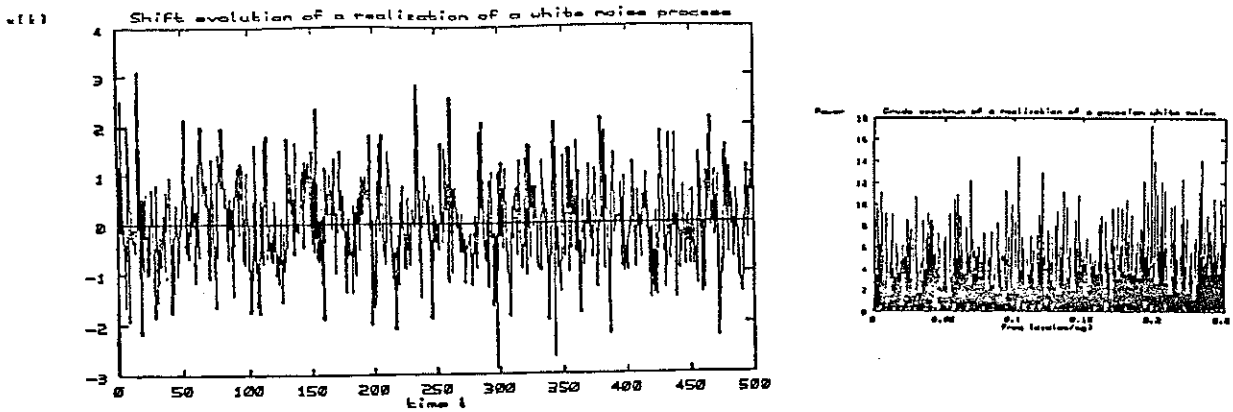


Figure C.6: *White random evolution.* Time evolution and crude power spectrum of a realization of a i.i.d. sequence of random variables with normal distribution N(0,1) (gaussian white noise).

C.3 Saw–Tooth Dynamics.

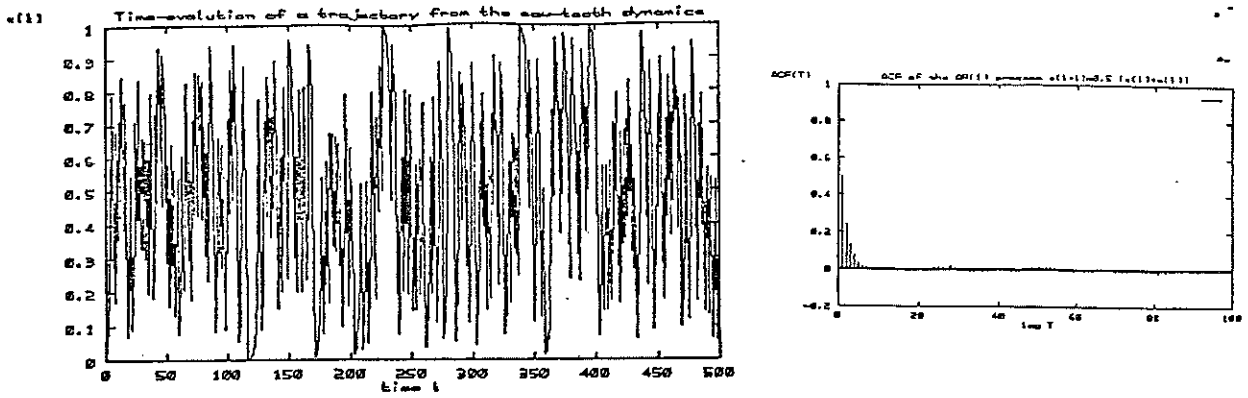


Figure C.7: *Generating a long orbit of the saw-tooth system.* Trajectory generated from the saw–tooth mapping (see **1.1.1**), that is $x_{k+1} = 2x_k - \text{int}(2x_k)$, $k = 1,\ldots,5000$. To avoid the numerical difficulties arising when a long orbit of this system is generated, we created a realization of length 5,000 of the AR(1) process $y_{k+1} = 0.5(y_k + \epsilon_{k+1})$, where $P(\epsilon_k = 0) = 0.5$ and $P(\epsilon_k = 1) = 0.5$ for all k, and we then inverted the sequence, i.e. we took $x_k = y_{-k}$, $1 \leq k \leq 5,000$. The autocorrelation function is also shown on the right.

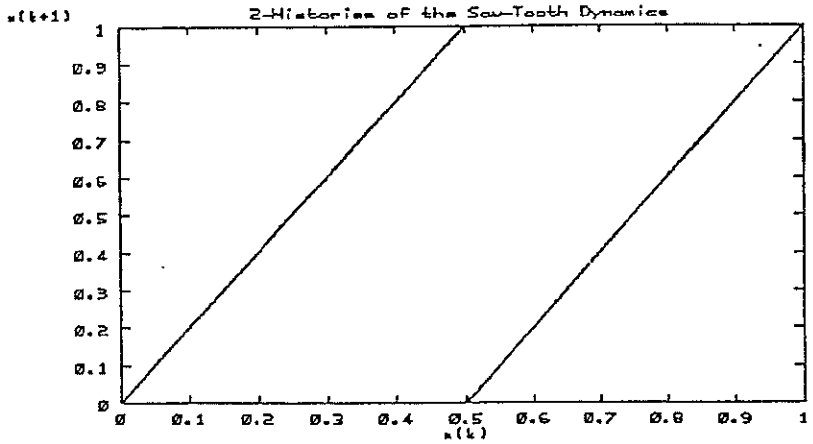


Figure C.8: *2d-Embedding of the saw-tooth dynamics.* 2-dimensional delay embedding of the univariate saw–tooth series in figure C.7.

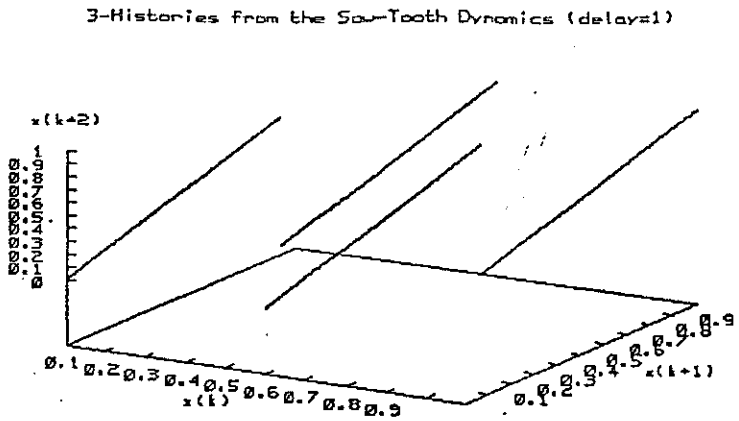
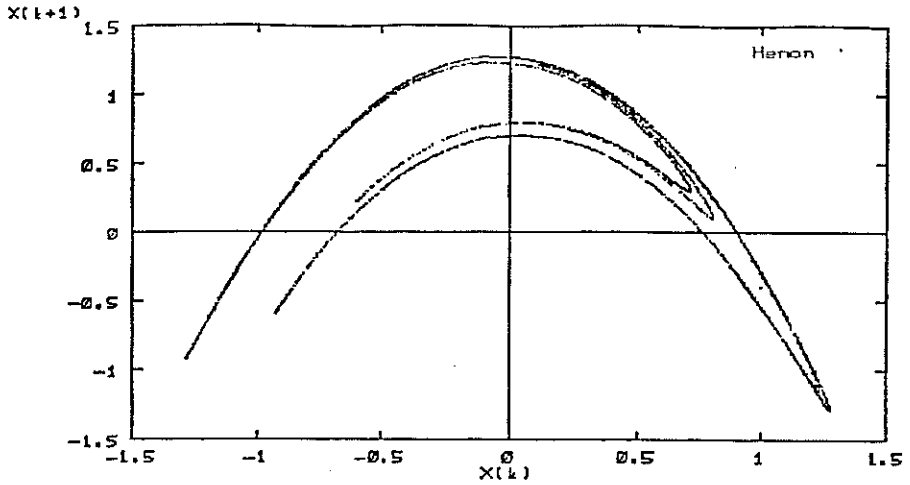


Figure C.9: *3d-Embedding of the saw-tooth dynamics.* 3d-Delay Embedding of the orbit of figure C.7. Notice that the 'phase' structure is not lost in dimension 3.

C.4 Delay Reconstruction of Strange Attractors.



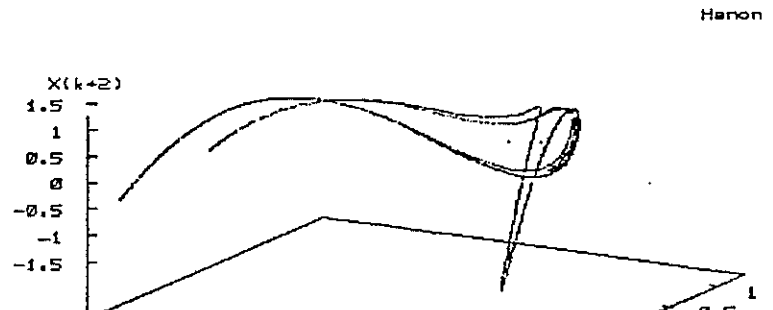
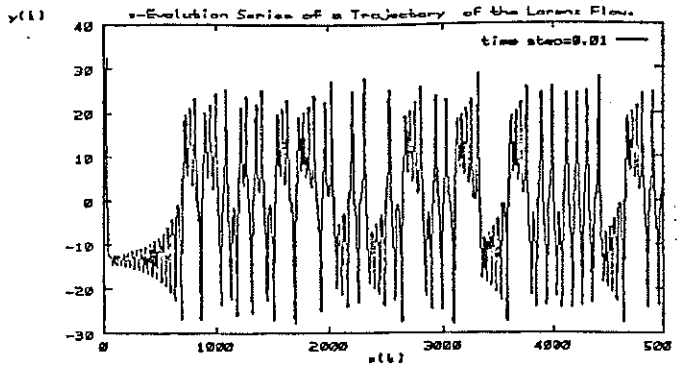
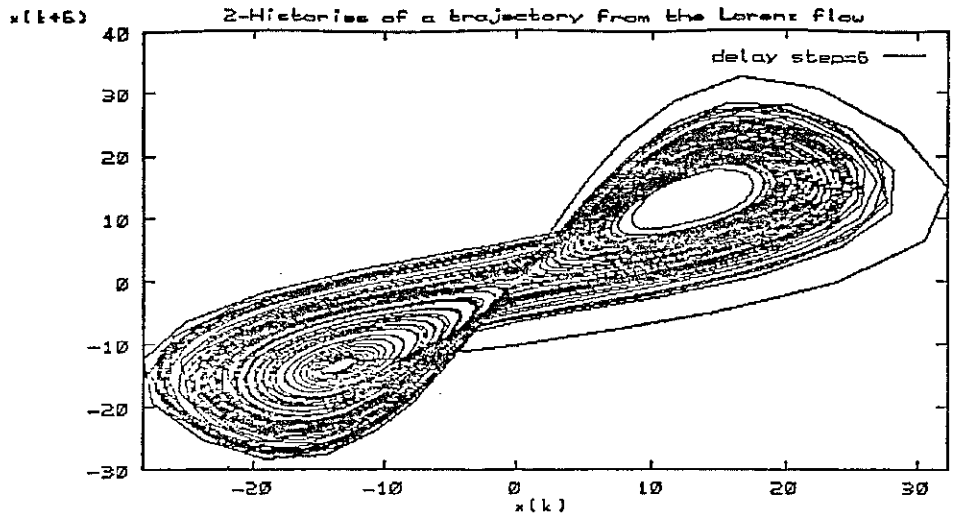


Figure C.10: *Delay embeddings of the Henon dynamics.* m-Histories, for $m = 2,3$, of a x–coordinate series monitored on a trajectory of the Henon system. This corresponds to $h(x,y) = \mathrm{pr}_1(x,y) = x$ in the Takens's theorem.





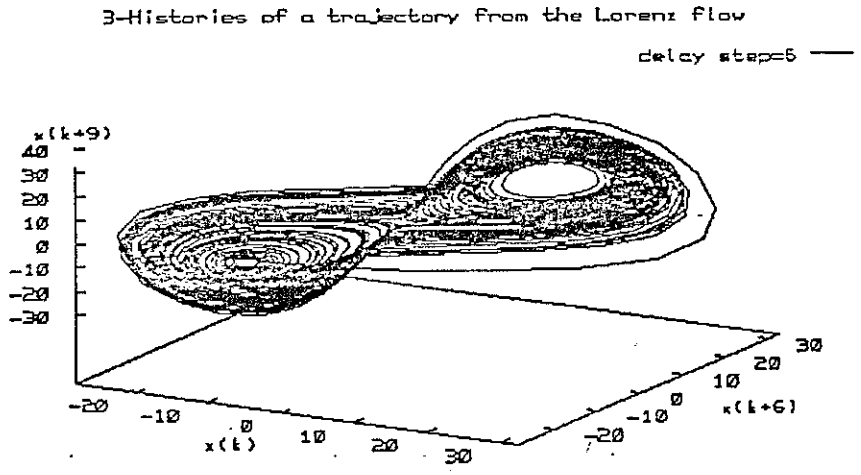


Figure C.11: *Delay reconstructions of the Lorenz dynamics.* m-Histories (m=1,2) created from a x–coordinate series observed in the trajectory of figure C.2 from the Lorenz system. The delay chosen was $\tau = 6$.

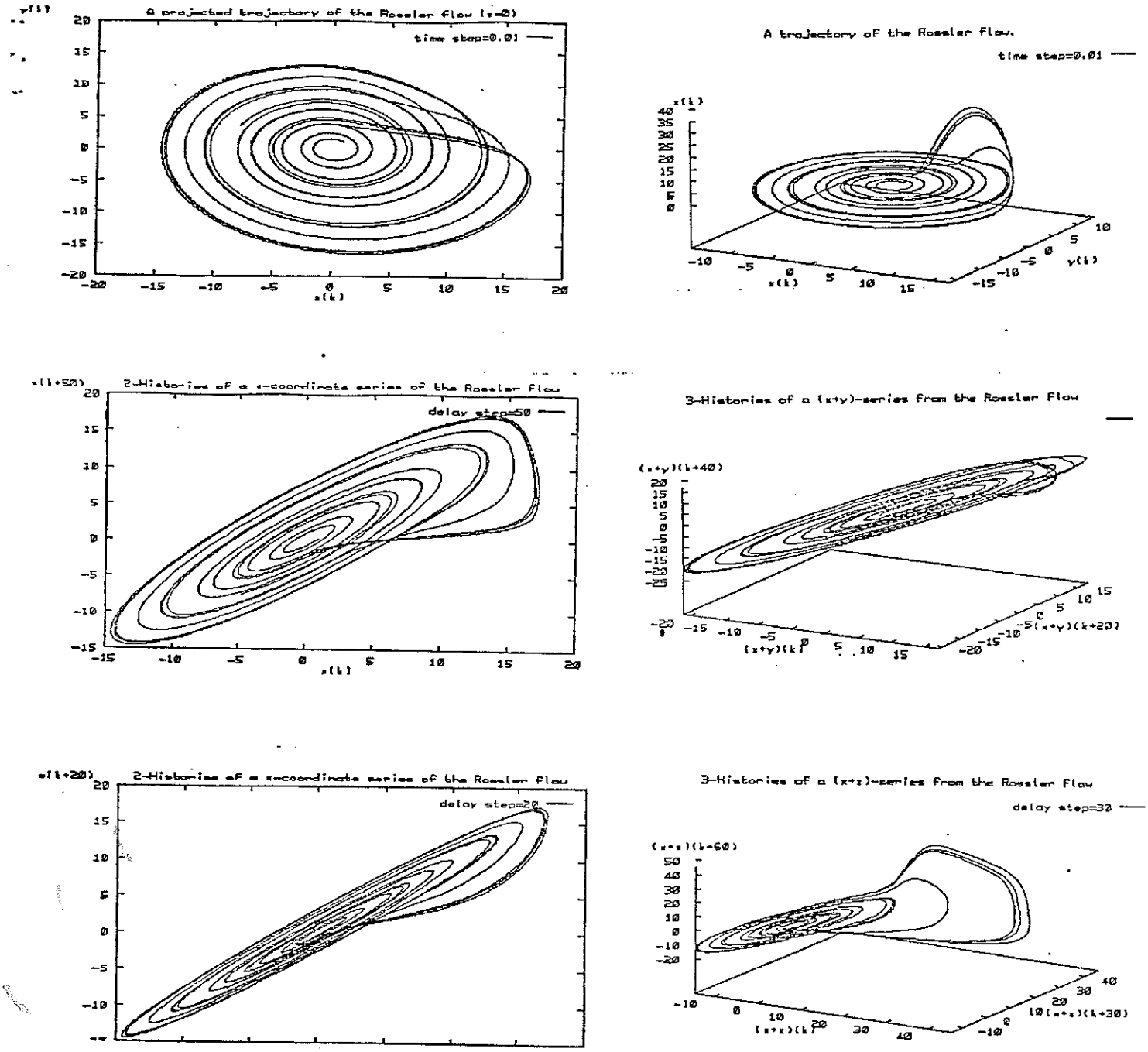


Figure C.12: *Delay Reconstructions of the Rössler dynamics.* The middle and bottom figures show different m-histories in dimensions 2 and 3, for different delay times, obtained from series recorded on a trajectory of the Rössler flow: the series of the x–coordinate of the trajectory, and the series of the sum of two coordinates (x+y or x+z) of the trajectory are used to create the m-histories of the dynamics (this corresponds to $h(x,y,z) = x + y$, $h(x,y,z) = x + z$ in the enmbedding theorems). These plots illustrate the embedding theorems of section 1.3 (compare this with the true dynamics, displayed at the top).

C.5 Correlation Dimension of the Processes of our Testing Laboratory.

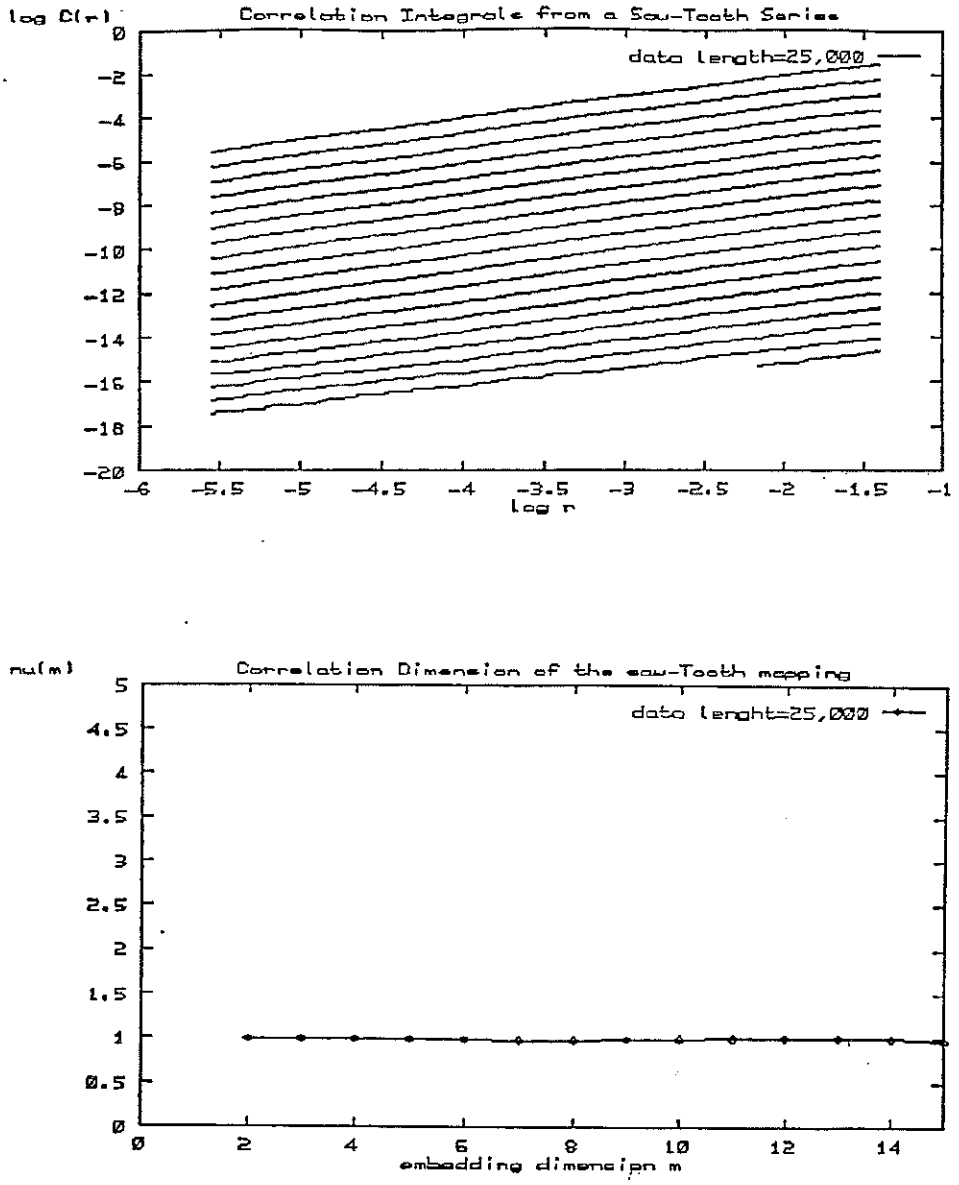


Figure C.13: *Correlation integrals of the saw-tooth dynamics.* The plots above show the correlation dimension analysis of the invariant measure associated with the dynamics of the saw–tooth mapping. The correlation integral of the time series (of length 25,000) of C.7 reveals the one-dimensionality of the Lebesgue measure (which is the physical invariant ergodic measure of the process). Notice that the estimate $\hat{\nu}(m) = 1$ is excellent up to dimension 16. From what was said in C.7 there follows the finite dimensionality of the stationary measure of the stochastic process AR(1) given there. (CPU time=19.66)

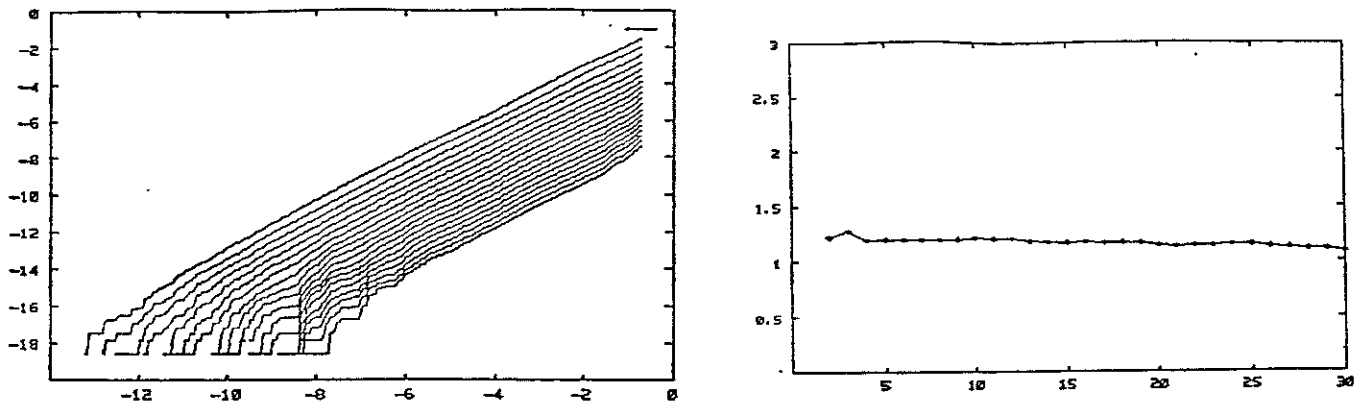


Figure C.14: *Correlation dimension of the Henon attractor.* Correlation integral plot from dimension $m = 2$ up to $m = 20$, and correlation dimension plot vs. embedding dimension m for the x–coordinate series of the Henon system of figure C.1. Notice the excellent agreement with the value $\nu \simeq 1.26$, accepted in the literature as the correlation dimension of the Henon attractor. The number of points of the orbit is 15,000. (CPU time = 7.68 min.)

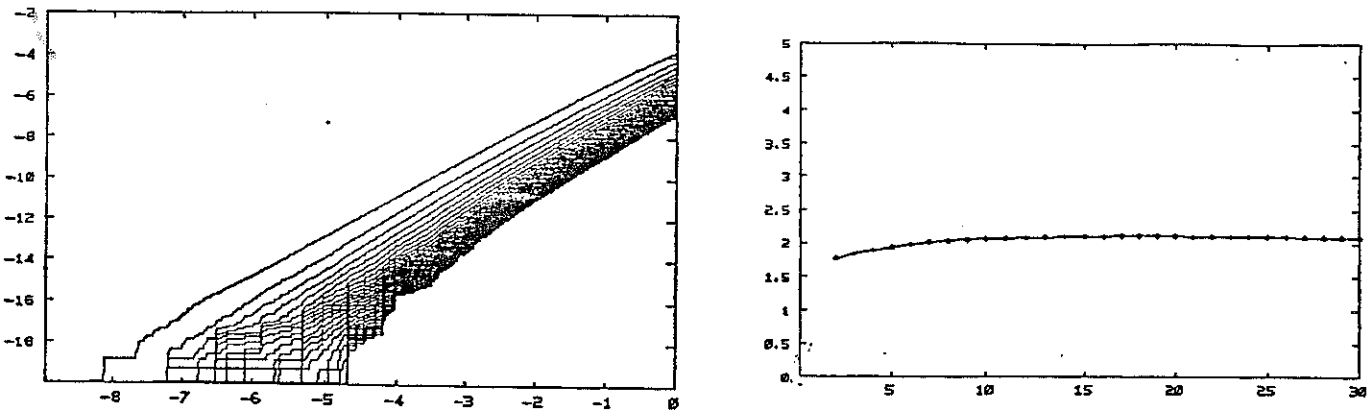


Figure C.15: *Correlation dimension of the Lorenz attractor.* The same figure as in C.14 for the x–coordinate series of the orbit in the Lorenz attractor represented in figure C.2. The value of the correlation dimension of the Lorenz attractor is $\simeq 2.05$. The embedding dimension runs from $m = 2$ to $m = 30$. The length of the input series is 30,000. (CPU time=4.53 min.)

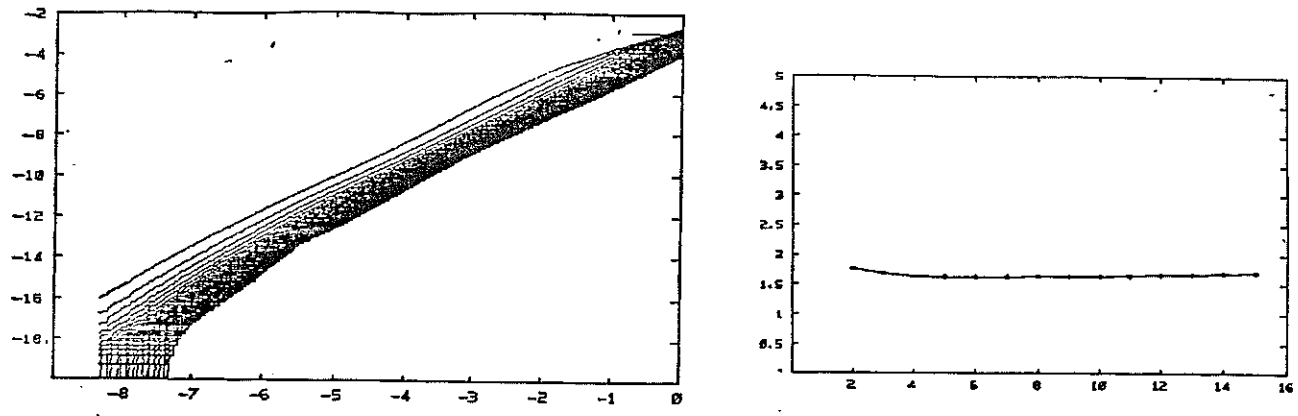


Figure C.16: *Correlation dimension of the Rössler attractor.* The same plots as those in figure C.15 for the observation series of the x–coordinate of the trajectory in C.3 of Rössler flow. The estimate value for the Liapunov dimension of the Rössler attractor is 2.01 [Russel et al.,80], which is in general larger than the correlation dimension. This is in accordance with our estimate. The length of the scanned series is 30,000. The dimension test (see chapter **2**) reveals the different nature of these series with respect to the series below (see figure C.17). (CPU time= 29.73 min.)

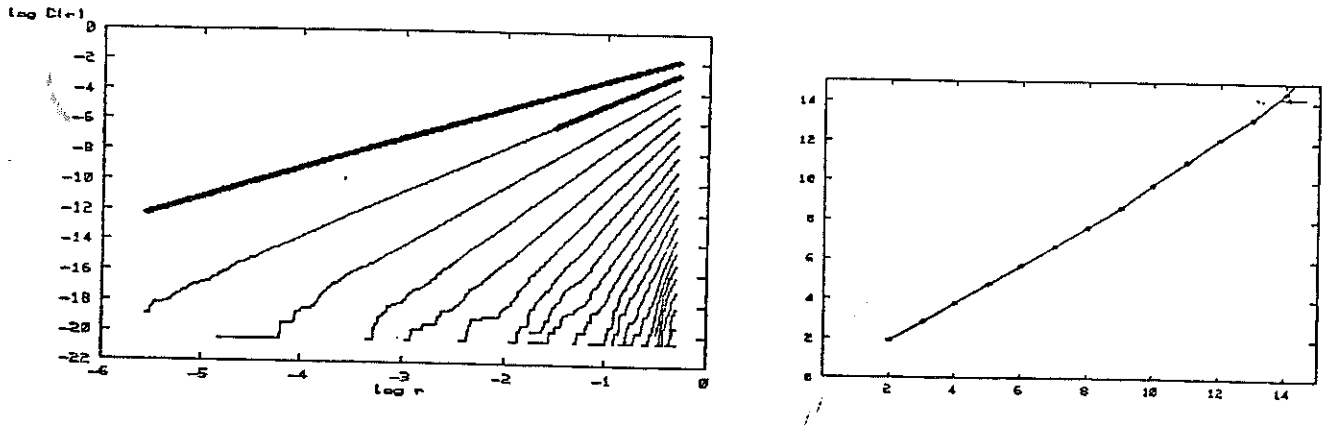


Figure C.17: *Correlation dimension of a space-filling noisy process.* Correlation integral plot and correlation dimension plot computed from a realization of length 30,000 of a stochastic process of random variables i.i.d. with distribution N(0,1) (see figure C.6). The correlation integral is plotted only up to $m = 20$ due to lack of data points (a slope cannot be reliably computed for dimensions near 20). The thicker superimposed lines were drawn from the points $(\log r_i, \log C(r_i))$, $i = 1, \ldots, 600$ for which the computation was performed. The space–filling nature of the process, in contrast to that of the systems above, is clear from the analysis. (CPU time=20.07 min.)

C.6 Robustness of Correlation Integral Estimates with respect to Smooth Observations.

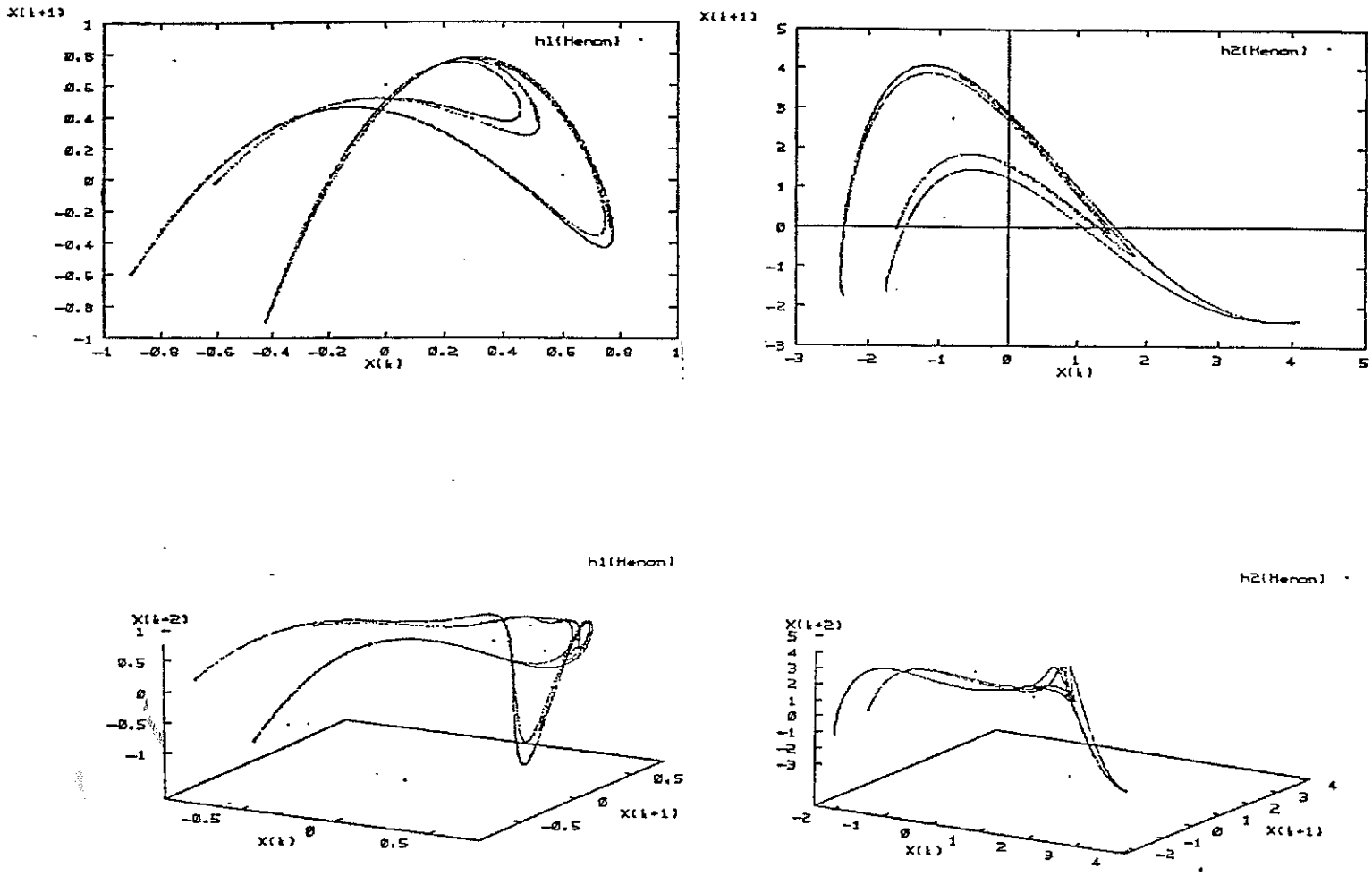


Figure C.18: *Smooth registrations on the Henon dynamics.* Two different series of length 25,000 were recorded on a trajectory of the Henon system by the smooth observables $h_1(x,y) = 0.3(2x+3y)$, and $h_2(x,y) = 0.8(-x+10y)$. The embedding reconstructions of the observed series $\{h_i(x_k,y_k)\}_k$, both for $i=1$ (left) and for $i=2$ (right), in dimensions 2 (top) and 3 (bottom) are displayed above. Notice that this corresponds to 'recording' functions in the Fractal Delay Embedding theorem (see section 1.3) distinct from the x–coordinate observation series (i.e. $h(x,y) = \mathrm{pr}_1(x,y)$) used so far; diffeomorphic images of the Henon attractor are obtained by the delay mapping in dimension three.

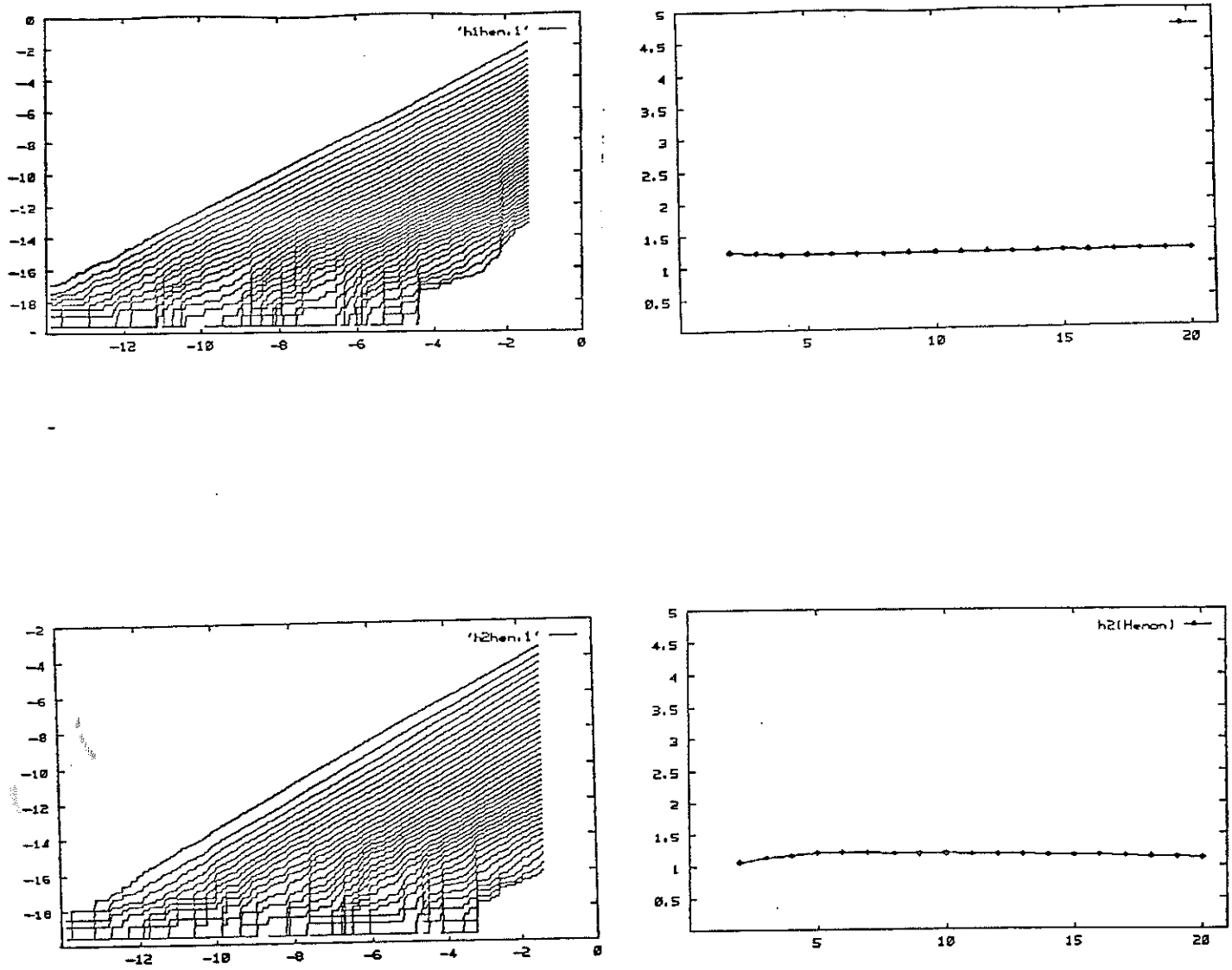


Figure C.19: *Correlation dimension of smooth recordings of the Henon dynamics.* Correlation integral and correlation dimension plots for the h_1-series (above) and for the h_2-series (below) from the figure C.18. Observe the excellent agreement with the results obtained for the genuine Henon attractor, thus illustrating the nice robustness of the dimension estimate. (CPU time=15.5)

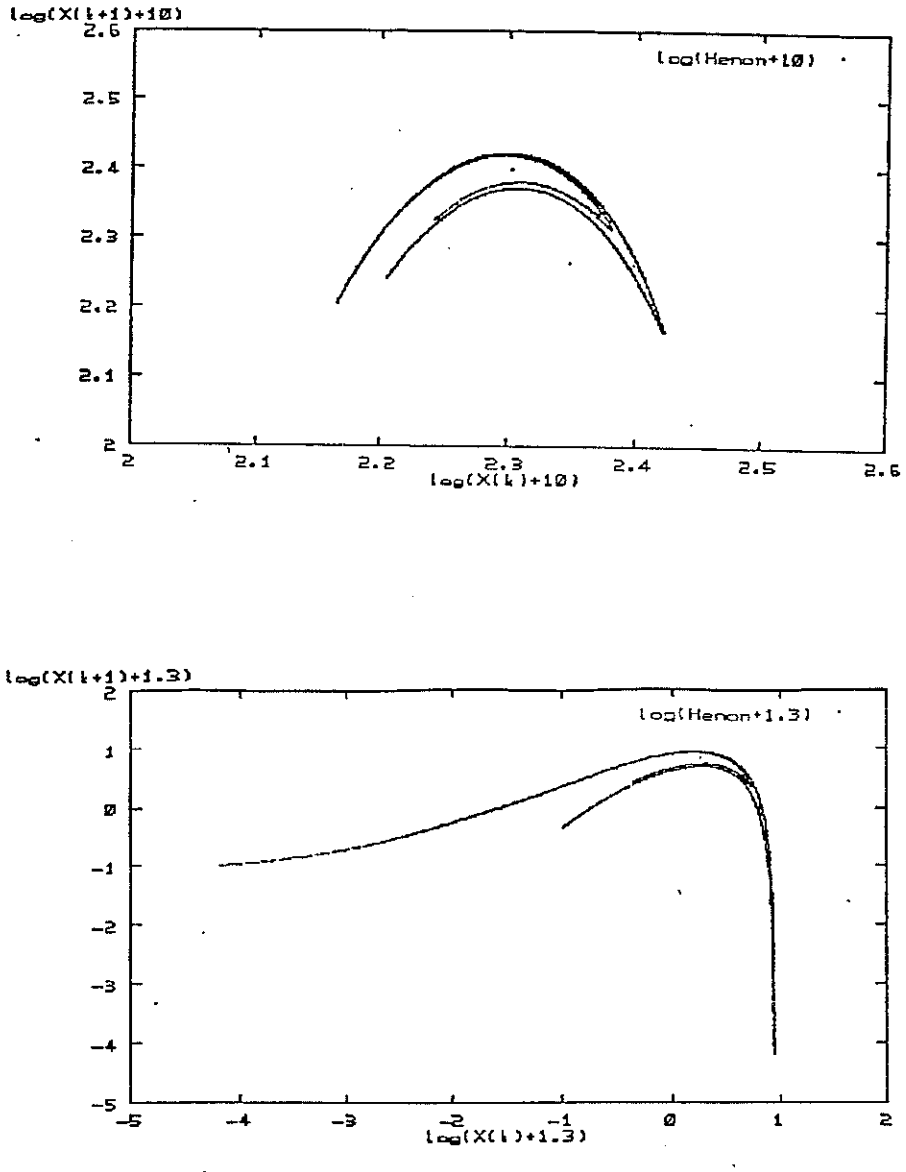


Figure C.20: *Logarithmic deformations of series recorded on the Henon system.* We considered two logarithmic series obtained from the x–coordinate sequence $\{x_k\}_{k=1,\ldots,15000}$ of the Henon system orbit in figure C.1. The figure shows the 2-histories of the series $\log_1 = \{\log(x_k + 10)\}_{k=1,\ldots,15000}$ (above) and $\log_2 = \{\log(x_k + 1.3)\}_{k=1,\ldots,15000}$ (below). Notice the scale in the plot at the top. This is due to the compressing effect of the logarithm in the series $\log_1$. This is only expected to influence the correlation integral at large scales.

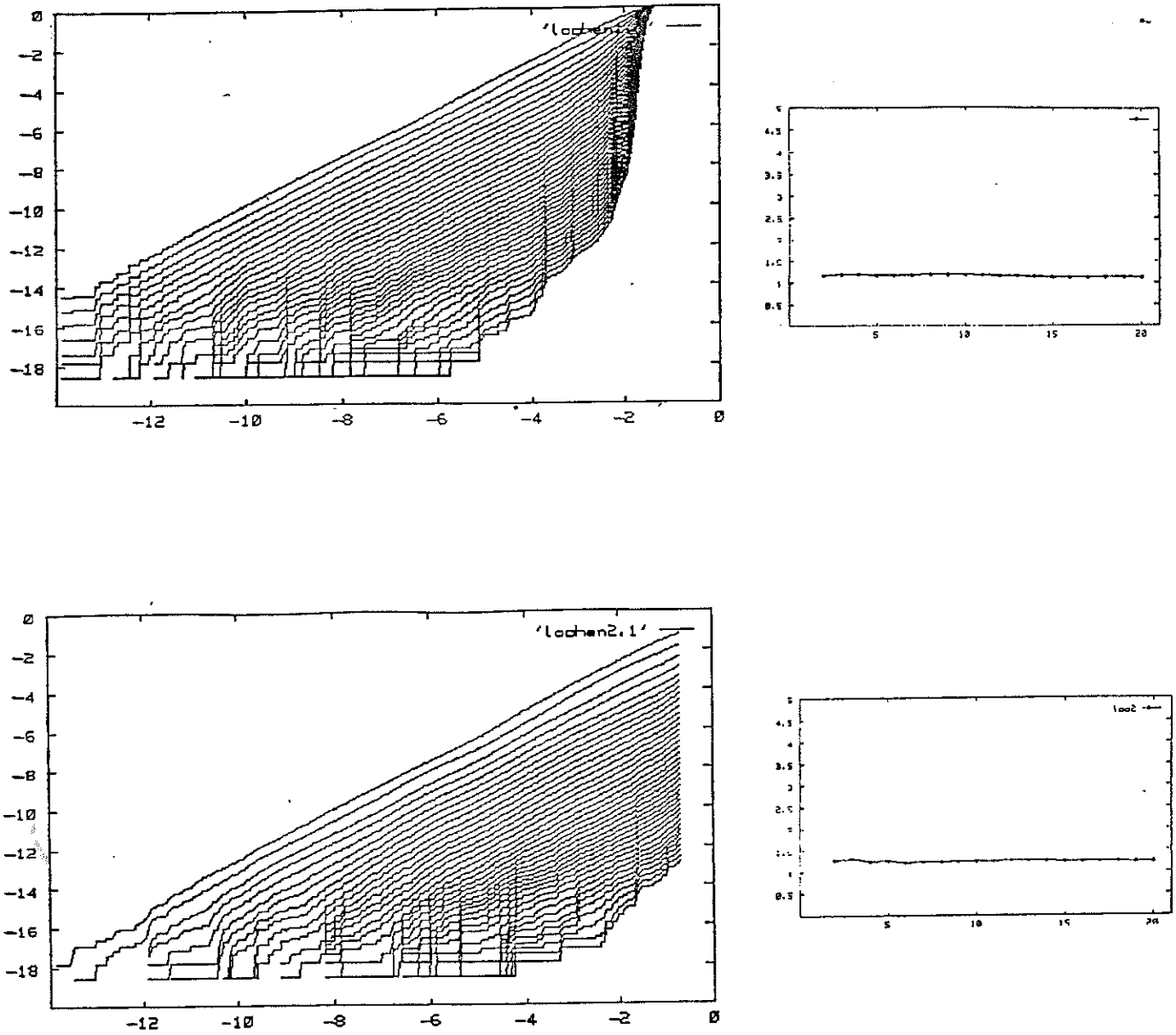


Figure C.21: *Correlation dimension of logarithmic series from the Henon dynamics.* Since we worked both with SWH-series and with their logarithmic series, we were interested in checking in practice the robustness/independence of numerical estimates of the correlation integral from a monitored signal when using, for the computation, finite–length logarithmically transformed series of the original one. For the sake of comparison, we computed the correlation integral of the series in figure C.20 under the parameter conditions of computations for the Henon attractor. As can be seen no difference is appreciated from the analysis above and that of C.14. The structure in the correlation plot of the series $\log_1$ (left side above) is due to the compressing effect mentioned in C.20. (CPU time =142.84 and 8.89)

C.7 Correlation Integrals of Attractors corrupted by Small–Amplitude Noises.

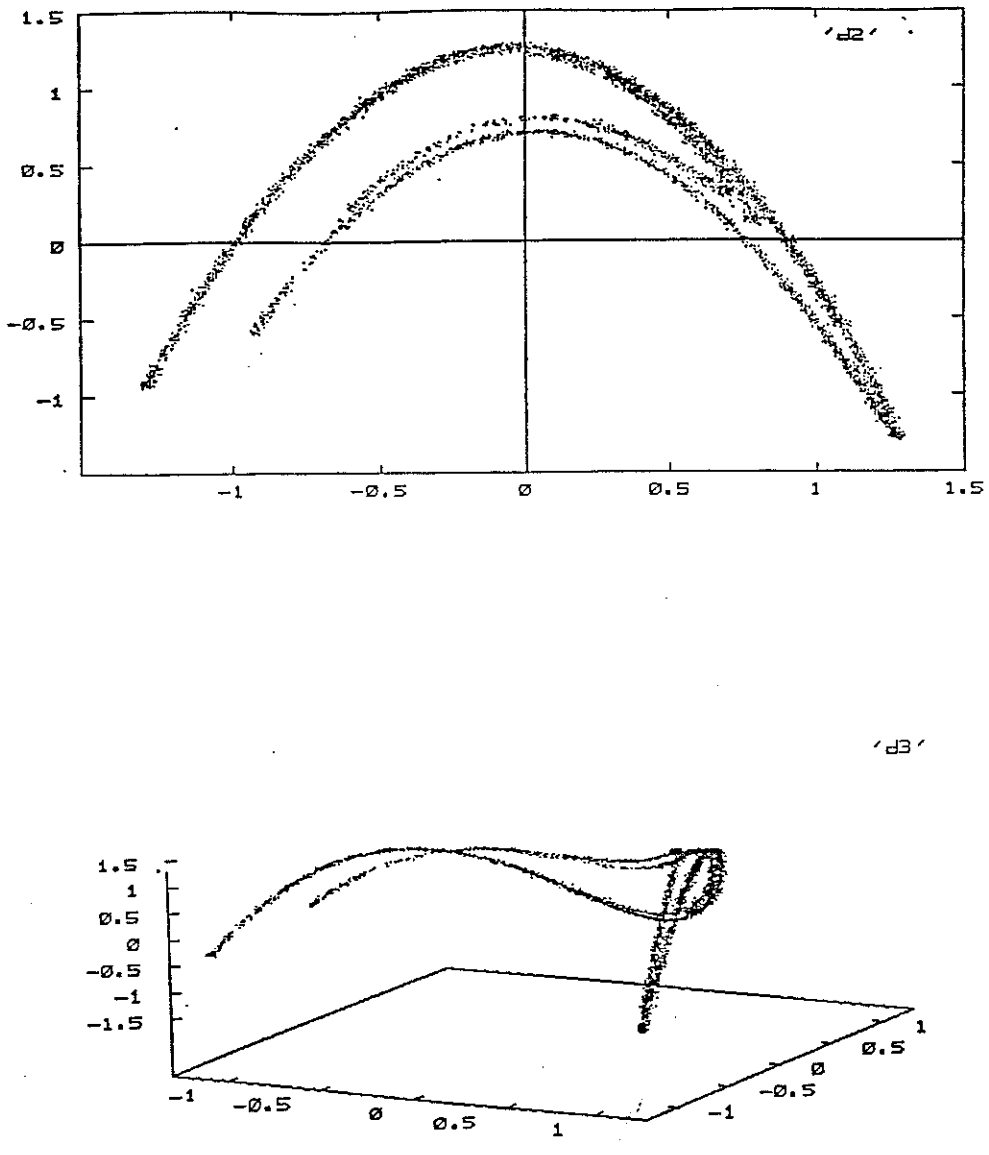


Figure C.22: *Delay reconstruction of a noisy chaotic orbit.* An orbit from the Henon system was corrupted by small–amplitude observational noise and the m-histories in dimensions 2 and 3 are shown above. The added noise was distributed according to a normal distribution N(0,0.01). Notice that the amplitude of the perturbation does not affect the large structure of the Henon attractor, whereas the space–filling introduced by the noise corrupts the attractor structure at small scales.

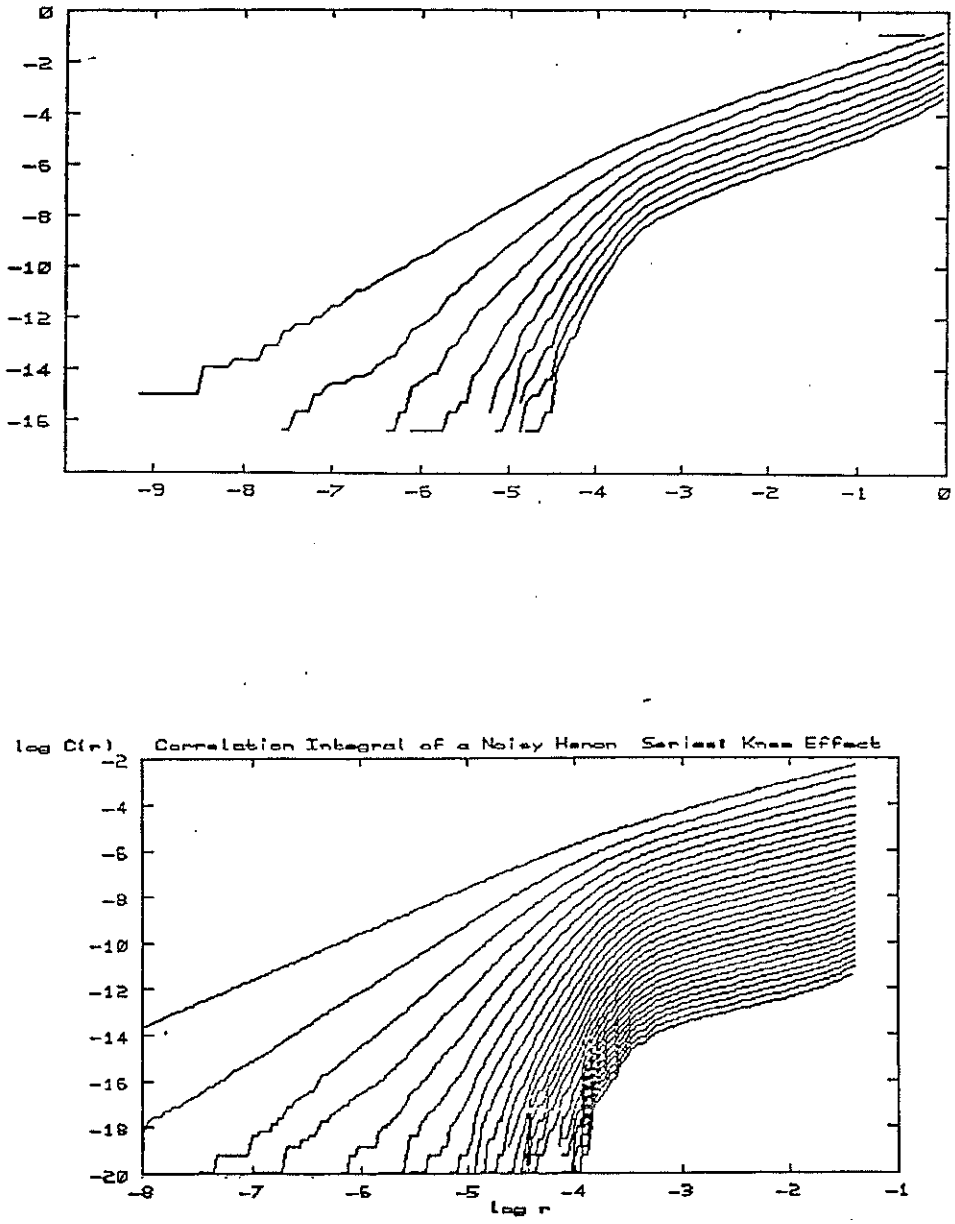


Figure C.23: *'Knee effect' in the correlation integral.* Correlation integral of the x–coordinate series of the noisy orbit of figure C.22. We considered 8,000 data points to go up to dimension 10 (top), and 30,000 data points to go up to dimension 40 (bottom). Notice that there are two distinct scaling regions, revealing the presence of two systems with different dimensionality acting at different scales. For resolution scales larger than the noise amplitude, the data set behaves like the Henon attractor; as soon as balls of small enough radii are considered, the mass spreading of the orbit is that of a space–filling process, that is, for all m $C^m(r) \simeq r^m$.

C.8 Correlation Integral Analysis of the SWH–Series.

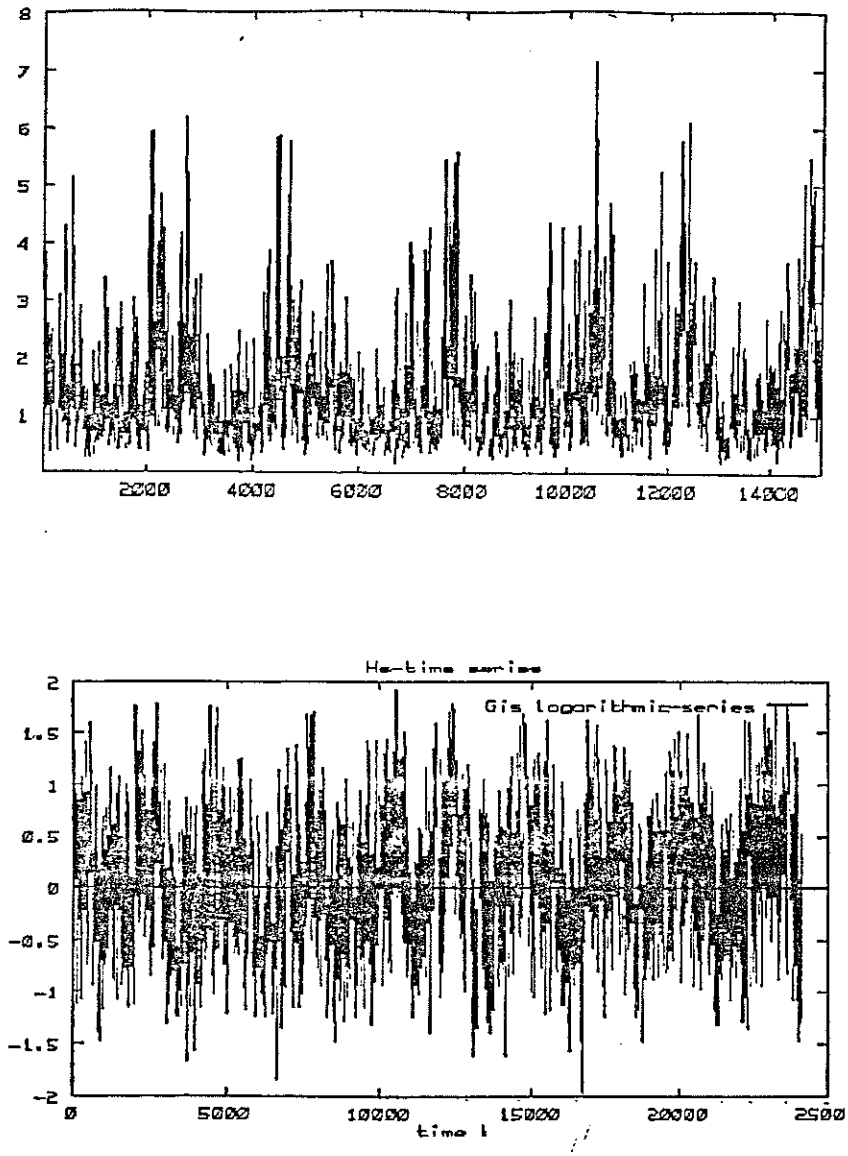


Figure C.24: *SWH-time series.* Time series monitoring significant wave height (SWH): at the top there appears the series Gi_{MO} (spectral method) from the first five years of data, and the bottom figure represents the logarithmic series $\log Gi_S$ (statistical method) of the whole recording period (10 years). See section **2.3.1** for the definitions.

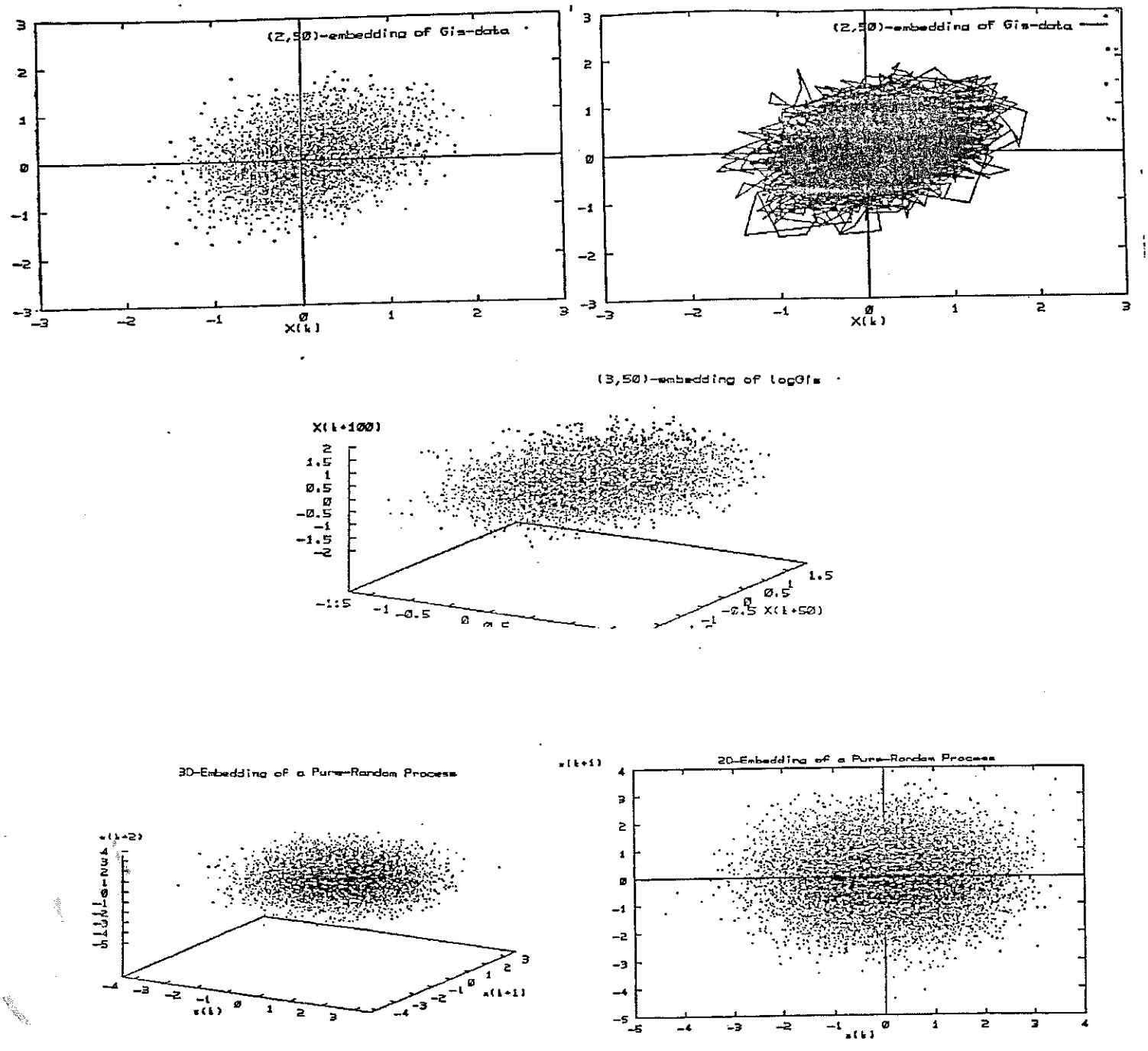


Figure C.25: *Delay plots of SWH-series.* Several delay reconstructions in dimension 2 and 3 of the Gis series (top) and $\log Gis$ series (middle). The reconstruction delay used above is $\tau = 50$, large enough to disentangle the spurious stretching along the lines $x = y$ (dimension 2) and $x = y = z$ (dimension 3). Our SWH–series have large correlation times. Thus the reconstructions above might not be optimal though they are significant enough to show that there is no structure in the m-histories of the series. Compare this with white–noise m-histories (bottom).

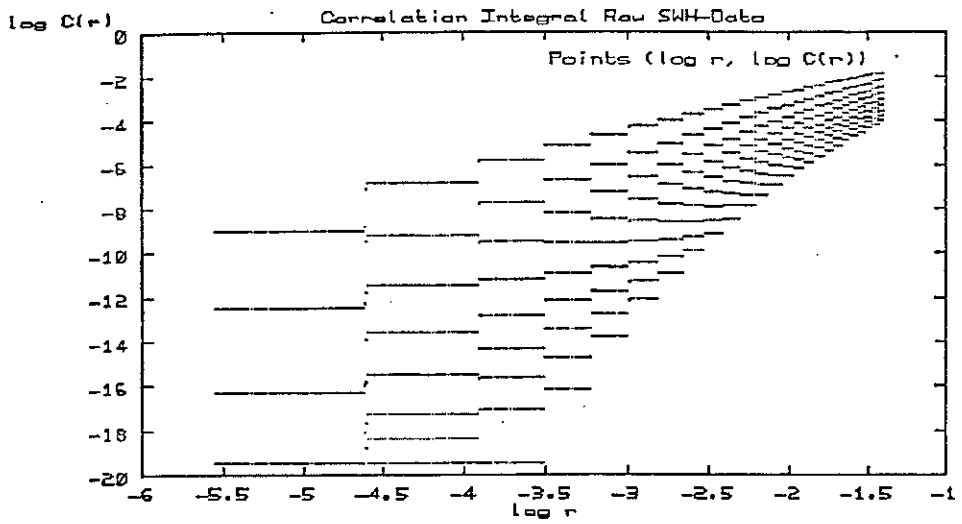


Figure C.26: *Crude correlation integrals of SWH-processes.* Correlation integral plots of H_S-series: 'stair-structure plots'. This figure shows the crude correlation integral plot of the series Gis, which exhibits the stepped-structure commented in section **2.3.2**. This is a problem to cope with when analyzing 'limited-observation data' (as H_S are), that might introduce a serious reduction of the number of different measurements (Gis has just 516 different data from the 24,098 recorded data). In chapter 2 we propose a method for analyzing these data in a significant way. See figure below.

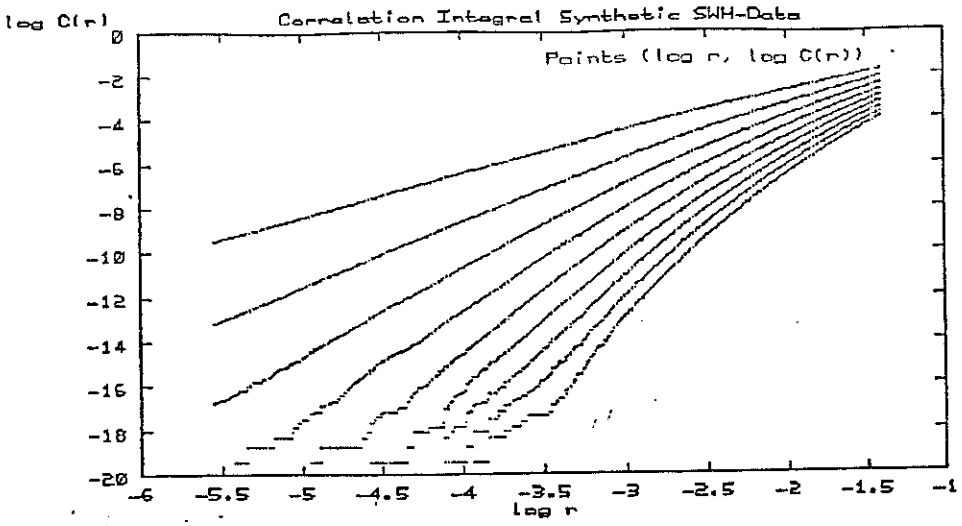


Figure C.27: *Analyzing 'limited-observation' data.* Ten different synthetic series were created by adding to Gis an i.i.d. process uniformly distributed on the interval $(-0.005, 0.005)$ (see section **2.3.3**). The correlation plots of one of these new series is shown above. Notice that the stair-shape in the plot of figure C.26 has dissapeared in the correlation integral plot of the new series. A knee expected to occur in the scale range $(0, 0.005)$ does not appear in the plot. This is a consequence of the space–filling nature of the SWH-processes. (CPU time=13.93 min.)

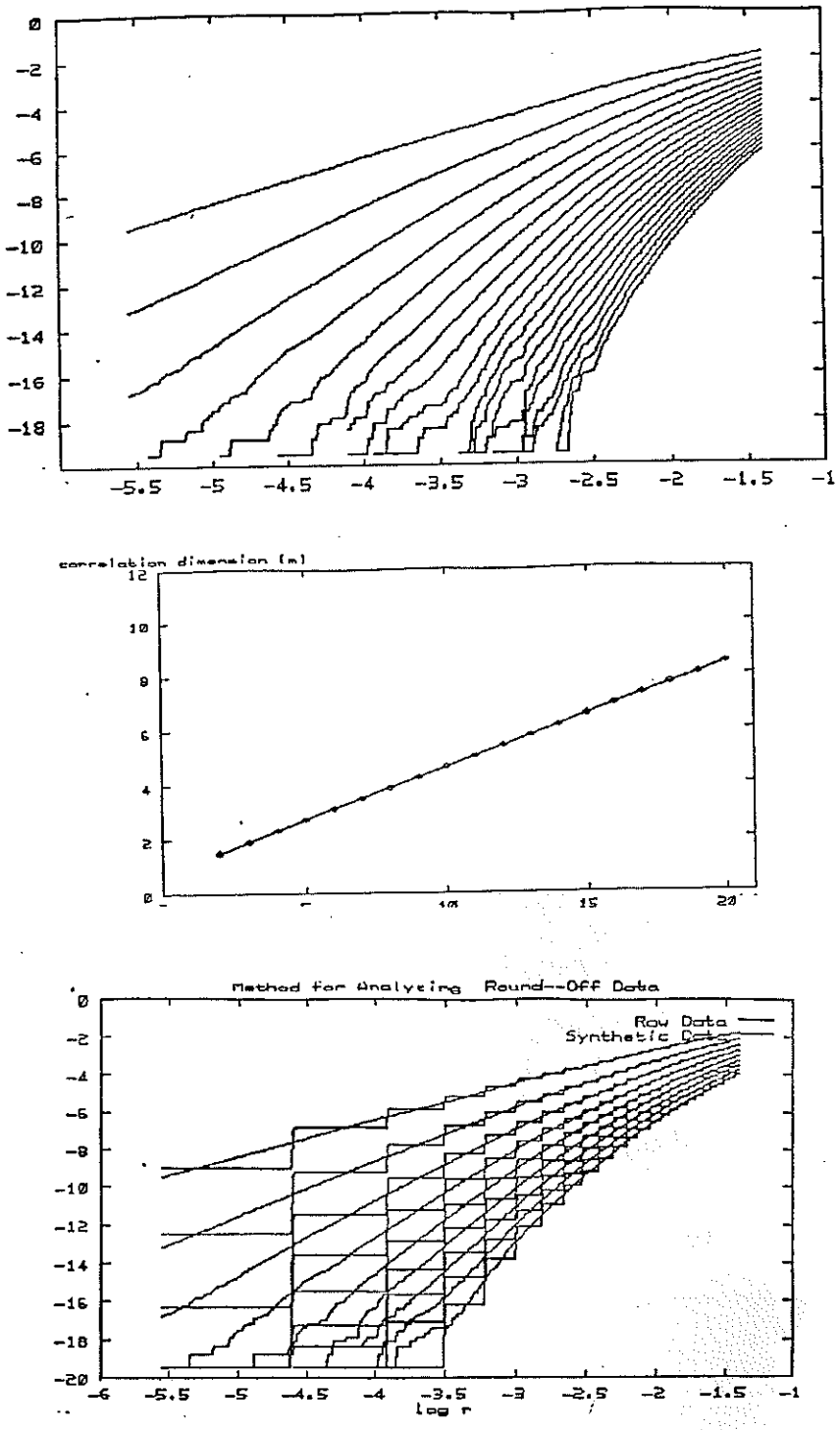


Figure C.28: *Correlation dimension of synthetic SWH-processes.* The correlation dimension plot obtained from an ensemble of ten series created from Gi_S by adding a small i.i.d. noise of amplitude 0.005 (see figure C.26) is shown above. The estimate $\nu(m)$ was obtained from the radii range outside the perturbation region. The analysis suggests a space-filling nature for the stationary measure of the H_S-process. The lower plot shows in the same graphic the original and the artificial correlation plots for the series $\log Gi_S$. This illustrates how the method works. (CPU time=127.64 min.)

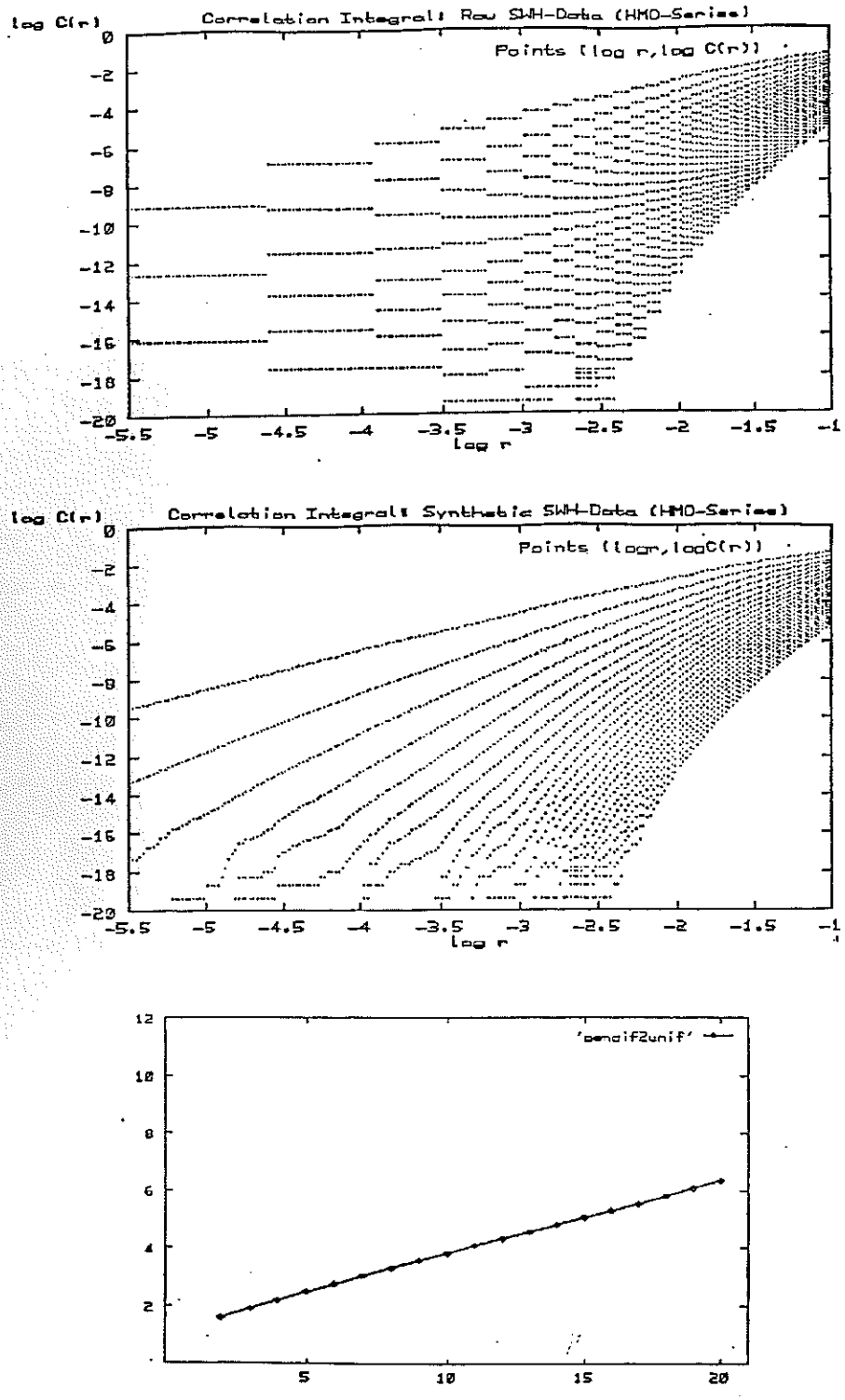


Figure C.29: *Correlation dimension of H_{MO}-processes.* The same as figures C.26 and C.28 for the SWH-series obtained from the spectral method, Gi_{MO}. There is no difference with respect to the analysis of the series Gi_S. The analysis of the logarithmic SWH-series rendered the same results as those reported here. (CPU time=119.79 min.)

C.9 K_2–Entropy of the Series from our Testing Laboratory.

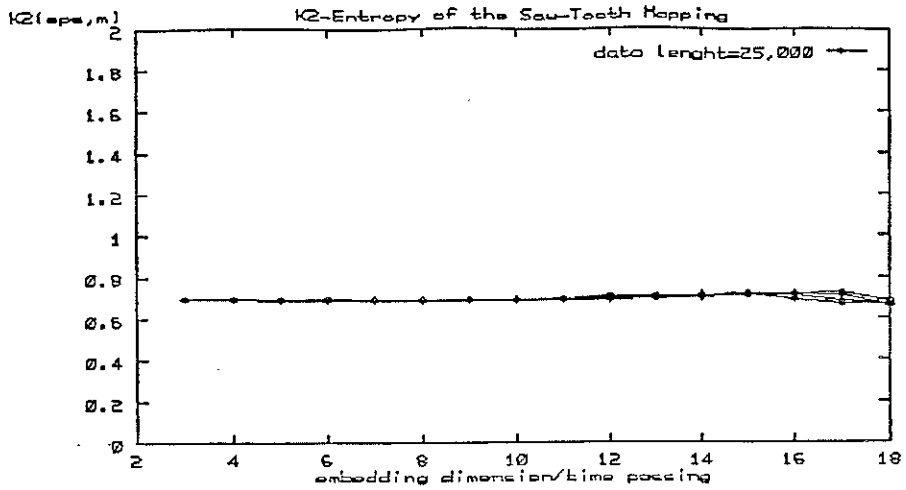


Figure C.30: *Entropy curves for the saw-tooth dynamics.* We plotted four entropy curves for the saw–tooth dynamics: $K_2(\varepsilon_i, m)$, $2 \leq m \leq 18$, and $\varepsilon_1 \simeq 0.15$, $\varepsilon_2 \simeq 0.09$, $\varepsilon_3 \simeq 0.06$ and $\varepsilon_4 \simeq 0.03$ (they are superimposed in the plot above). Notice that the K_2-entropy of the saw–tooth dynamics (with the Lebesgue measure) is theoretically known, namely $K_2 = \log 2 \simeq 0.693$. Compare this with the estimated value above.

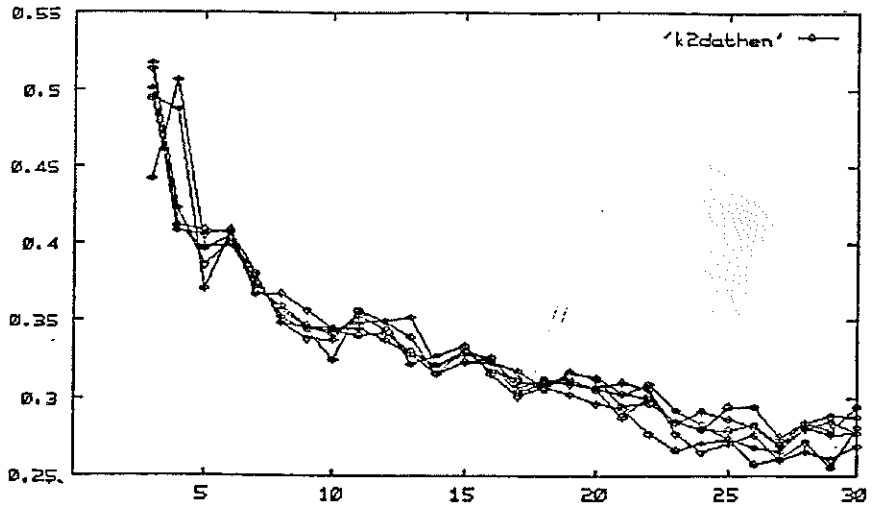


Figure C.31: *Entropy curves of the Henon mapping.* We have plotted six entropy curves $K_2(\varepsilon, m)$ up to $m = 20$, and for radii values varying from $\varepsilon_1 \simeq 0.4$ to $\varepsilon_6 \simeq 0.005$. The curves converge to a well-defined value that provides an estimate from below of the positive Liapunov exponent of the system $\lambda_1 \simeq 0.42$ (see chapter 4).

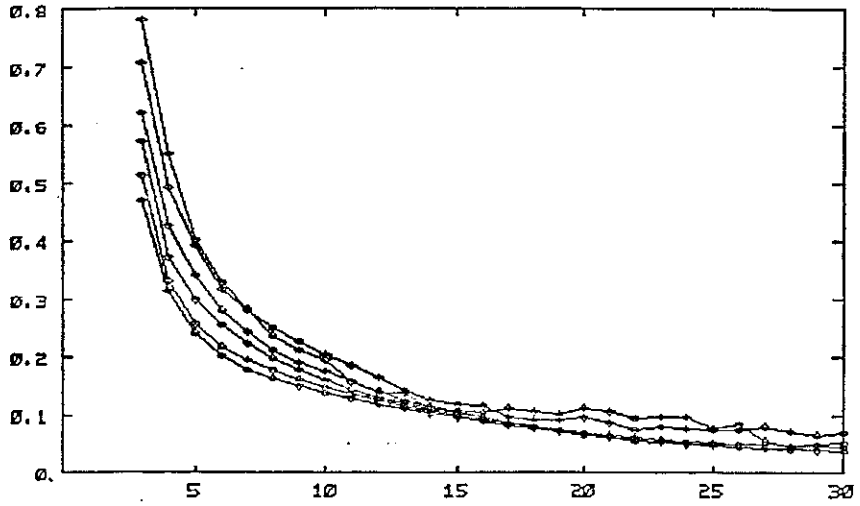


Figure C.32: *Entropy curves of the Lorenz flow.* We have plotted six entropy curves $K_2(\varepsilon, m)$ up to $m = 30$ and $\varepsilon \in \{0.5, 0.4, 0.15, 0.08, 0.035, 0.02\}$. A nice convergence $K_2(\varepsilon, m) \mapsto K_2 \simeq 0.05$ as m increases (independently of ε) is appreciated from the plot above, in accordance with the expected value.

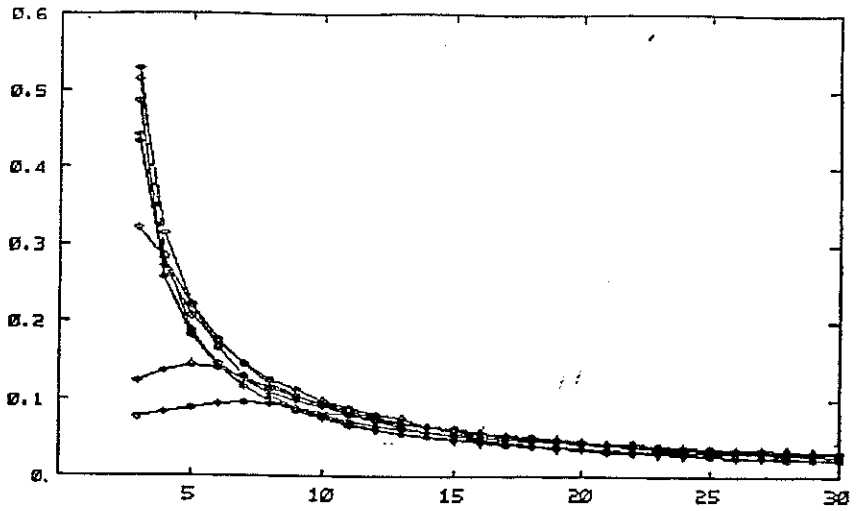


Figure C.33: *Entropy curves for the Rössler flow.* The same curves as in as figure C.32 are drawn above, for the Rössler system. The values of ε's are quite similar to those above. Notice the convergence of the entropy curves to a well defined finite and positive value, which is a consequence of the stochasticity (caused by the SIC) of the system.

C.10 Behaviour of the K_2–Entropy under Smooth Transformations.

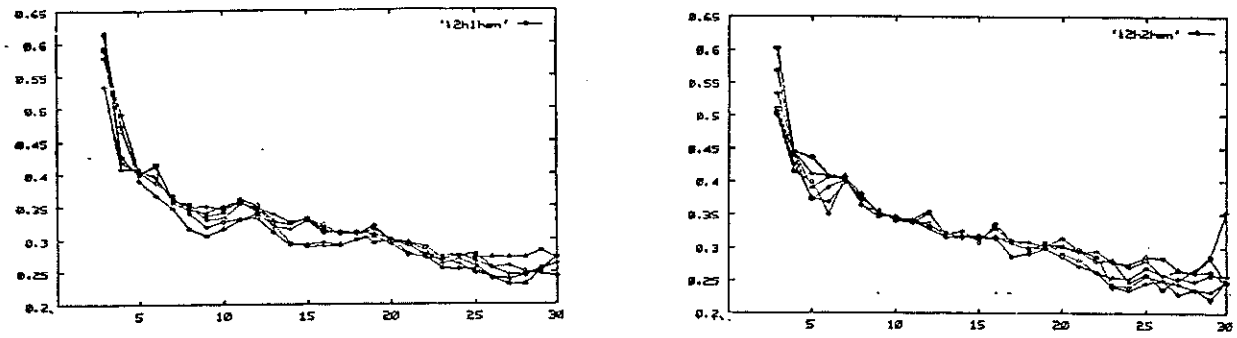


Figure C.34: *Entropy curves of transformed series from the Henon dynamics.* The plot above shows the entropy curves $K_2(\varepsilon, m)$, corresponding to the same ε's as in C.31, for the transformed series observed in the Henon attractor given in C.18. Notice the nice robustness of the value estimated as the K_2–entropy, thus showing the invariance of this statistic under smooth mappings.

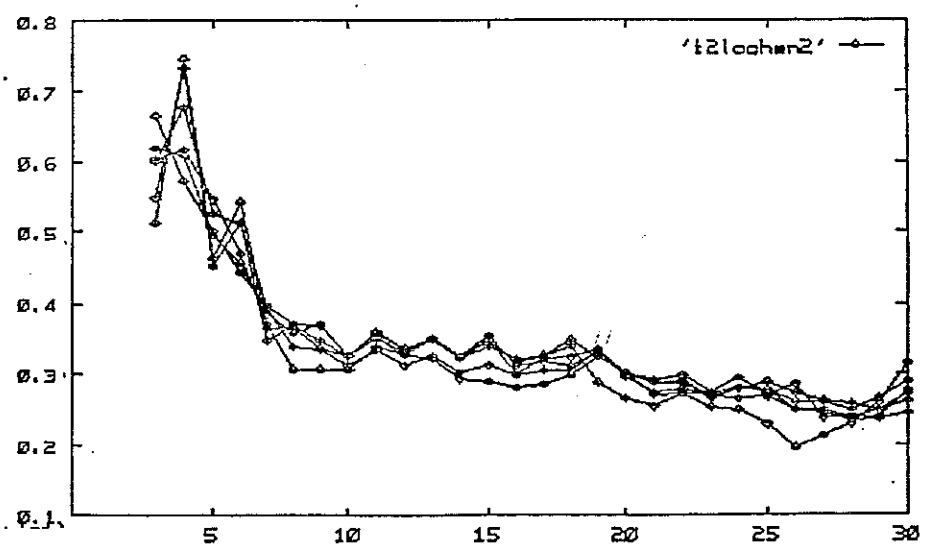


Figure C.35: *Entropy of a log-transformed x-series from the Henon dynamics.* This plot shows the invariance of entropy under logarithmic transformations. Specifically, the series from C.20, $u_k = \log(x_k + 1.3)$ with x_k an x–coordinate series from the Henon system, is considered; six entropy curves corresponding to those computed for the Henon x-series (see C.31) are calculated.

C.11 Reading Features of the System Randomness from the Entropy Curves.

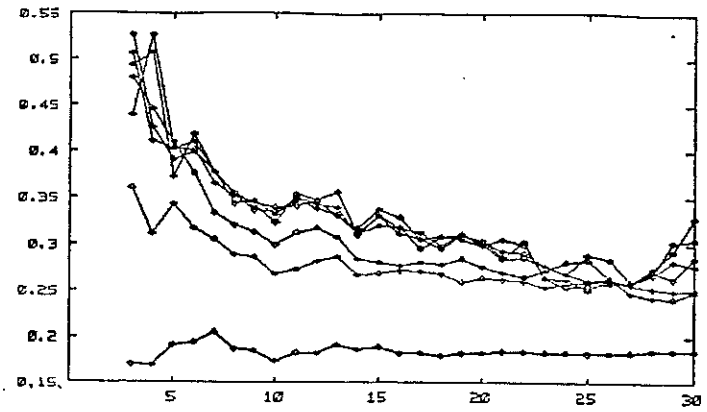


Figure C.36: *Entropy curves of logarithmic series from the Henon system.* We use the Entropy curves of the logarithmic series $u_k = \log(x_k + 10)$, where $\{x_k\}_k$ is the x-series of C.1 from the Henon dynamics, to illustrate the dependence of the entropy on the size ε of the partition of the phase space. The lowest curve above corresponds to the curve $K_2(0.15, m)$, which does not converge to the true value of the entropy for the Henon system as the other five entropy curves do converge. This is due to the compressing effect of the logarithm on the Henon series (see figure C.20), which implies in this case that the system has as much randomness (measured by the K_2-entropy) as $K_2 \simeq 0.18$ when the system dynamics in the phase space is seen at a resolution scale of $\varepsilon \simeq 0.15$. The true value of $K_2 \simeq 0.3$ appears as soon as the smaller resolutions are considered.

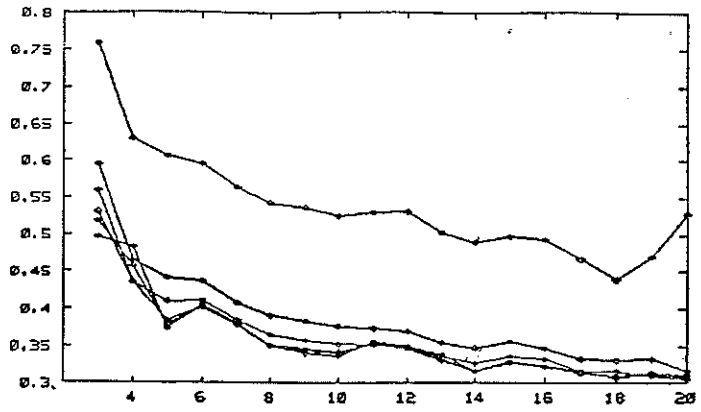


Figure C.37: *Entropy curves of chaotic orbits observed with small-amplitude noise.* This is the counterpart of the 'knee effect' for the correlation dimension. The noise becomes important at small scales so that the entropy curves of small radii give higher estimates of the entropy. This is shown above for the series of the x–coordinate in C.22, which had an added noise of amplitude 0.01: for the five 'larger' values of $0.09 \leq \varepsilon \leq 0.03$ the entropy curves converge to the entropy of the Henon system $K_2 \simeq 0.3$, but for $\varepsilon_6 \simeq 0.01$ the entropy curve $K_2(0.01, m)$ (upper curve) does no longer converge to this value, thus confirming the increase of the stochasticity in the system within the scales where the space–filling noise dominates. See C.39 for the entropy plot of a space-filling process.

C.12 Entropy of SWH–Processes.

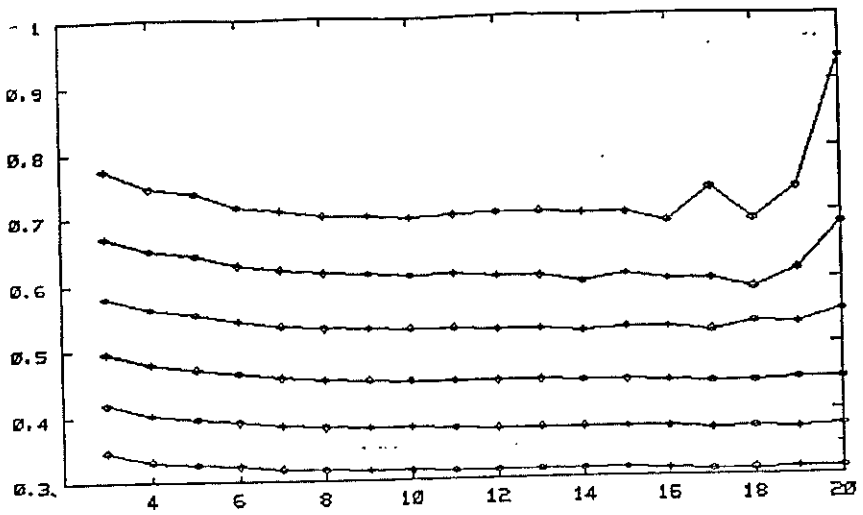


Figure C.38: *Entropy curves of H_S-processes.* No convergence of the entropy curves is apparent from the plot above, in contrast to the examples in the preceding section. The lower curve corresponds to $K_2(0.2, m)$ and the upper entropy curve is $K_2(0.08, m)$. The upper–shift of the $K_2(\varepsilon, m)$ curves as ε decreases reflects the increase of randomness in the time evolution defined by the shift mapping in the SWH-process. This is consistent with pure random behaviour, see figure below. The entropy analysis of the logarithmic series gives the same results.

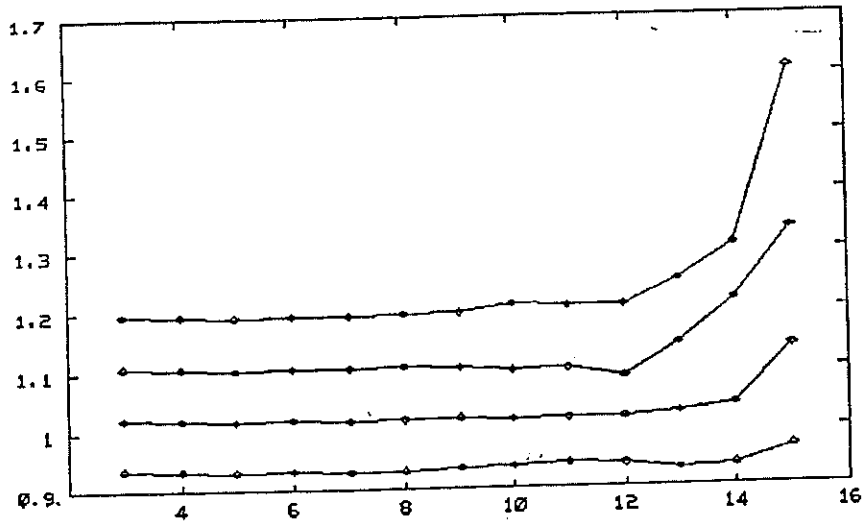


Figure C.39: *Entropy curves of a white (gaussian) noise process.* The figure shows the characteristic behaviour of pure random processes in terms of the K_2-entropy: the 'non-convergence' of the entropy curves $K_2(\varepsilon, m)$ as the size of the partition decreases. The noise considered is that of figure C.17. The lack of data (just 30,000 data) does not permit a significant computation up to a high dimension considering at the same time small values for ε. In fact, the radius considered above goes just from $\varepsilon_1 \simeq 0.7$ to $\varepsilon_4 \simeq 0.5$. This lack of data provokes the spurious increase at values of $m \geq 12$ (notice that this effect worsens for smaller radii. The behaviour sketched above is esentially that of the SWH-series (see figure above).

C.13 Correlation and Entropy Analysis of Filtered SWH-Series.

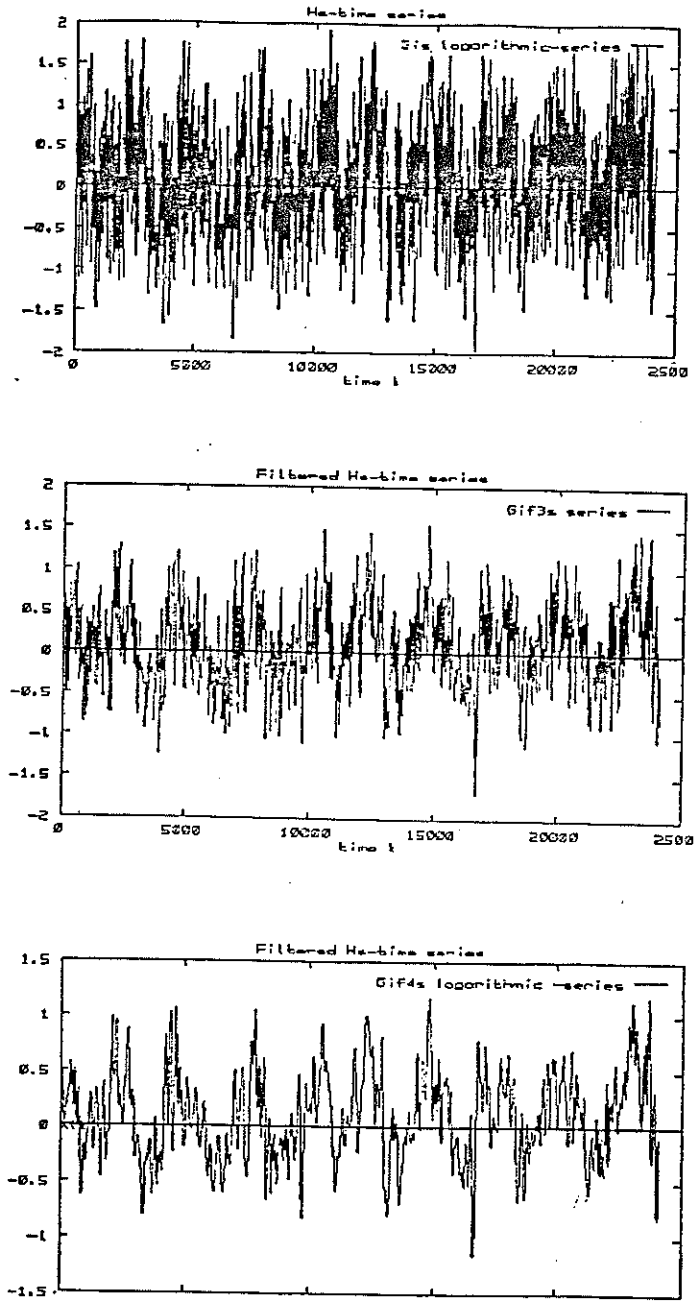
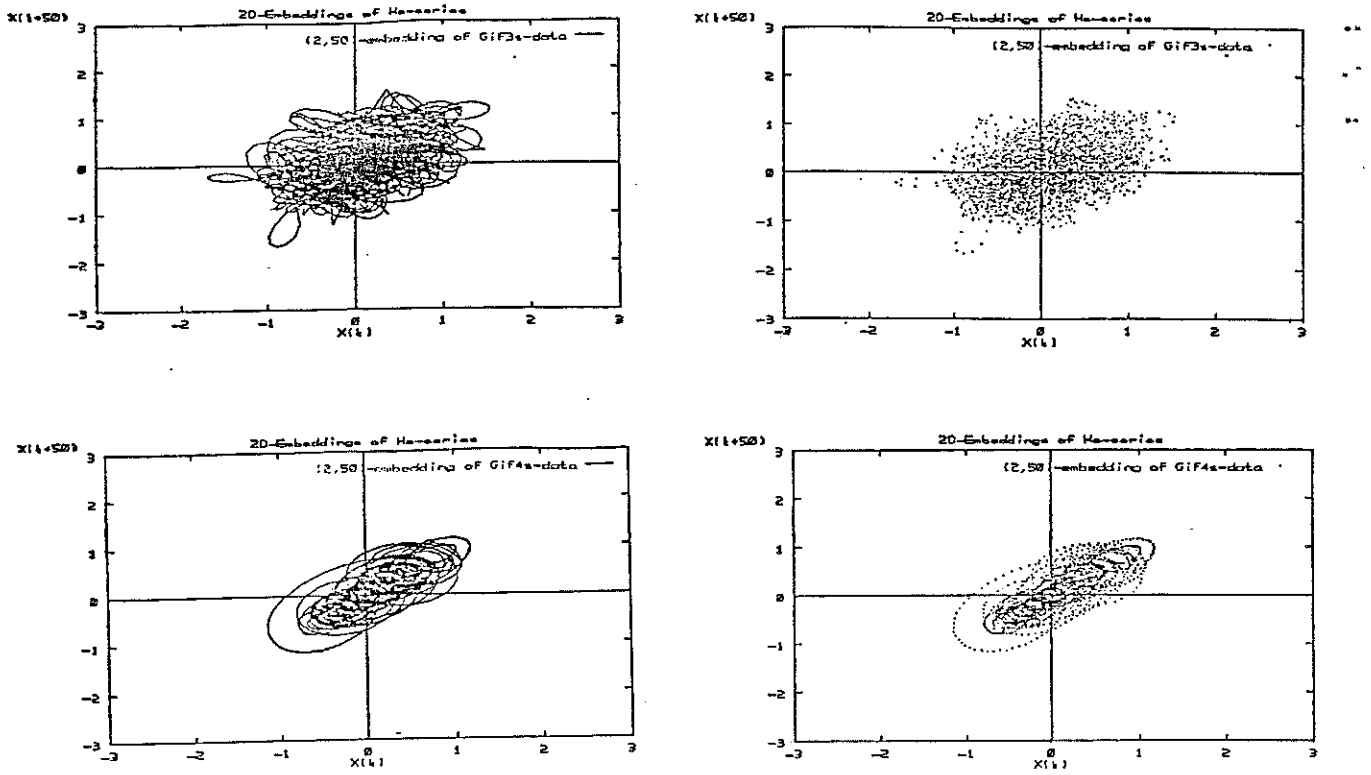


Figure C.40: *Filtering SWH-series.* The original SWH-series logGi_S (top), and the output series after imposing low-pass filtering: log$Gif3_S$ (middle) and log$Gif4_S$ (bottom). See section 3.2.2 for the frequency cut-off of each series. Notice that the time evolution of the f4-filtered series is far from being trivial, as the spectral analysis shows.



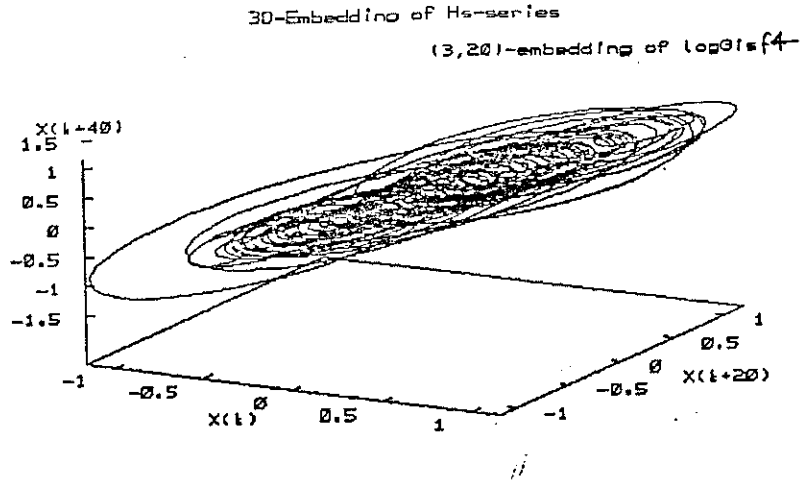
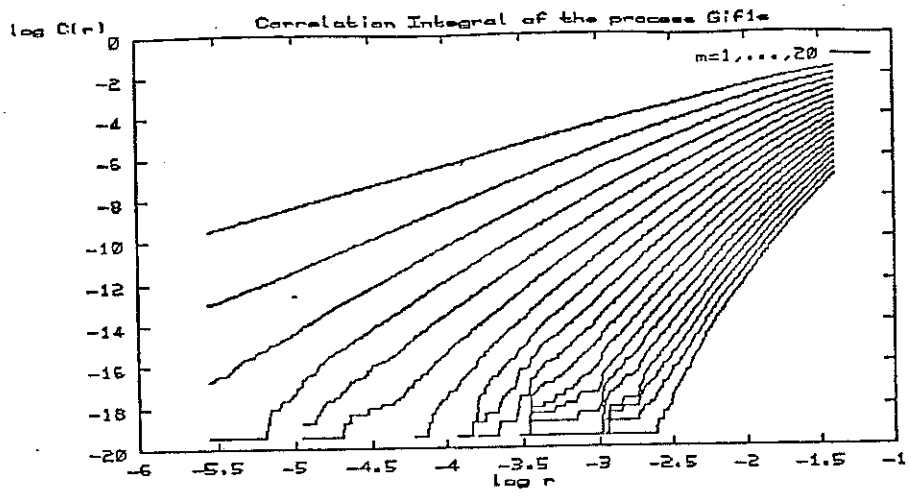


Figure C.41: *m-Histories (m=1,2) of the filtered SWH-series.* This figure shows several delay reconstructions of the filtered (logarithmic) series in figure C.40. The delay steps used for the construction of the m-histories are large ($\tau = 20, 30, 50$) but far below the delay suggested by the decorrelation time criterion (see section **1.3.**).



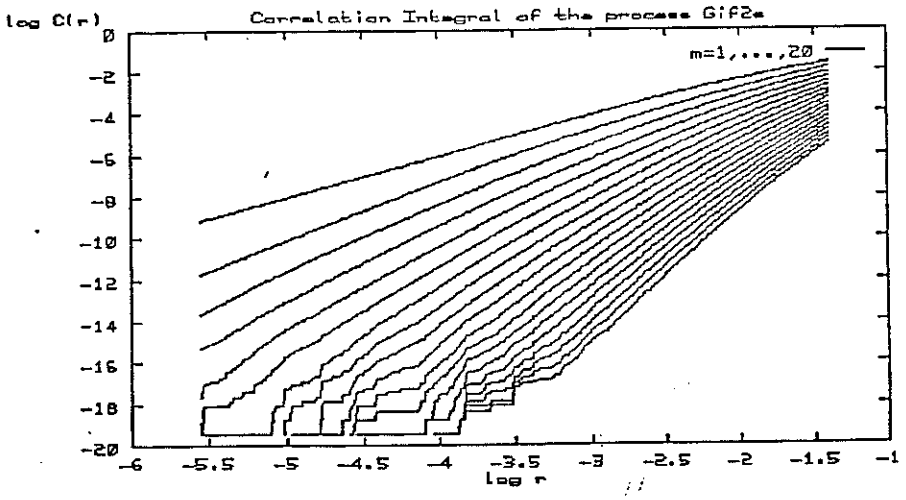


Figure C.42: *Correlation integrals of the filtered SWH-series $Gif1_S$ and $Gif2_S$.* Notice that there is no qualitative difference between this and the correlation plots of the original SWH-series Gi_S (see figure C.26. (CPU times= 17.77 and 23.81 min.)

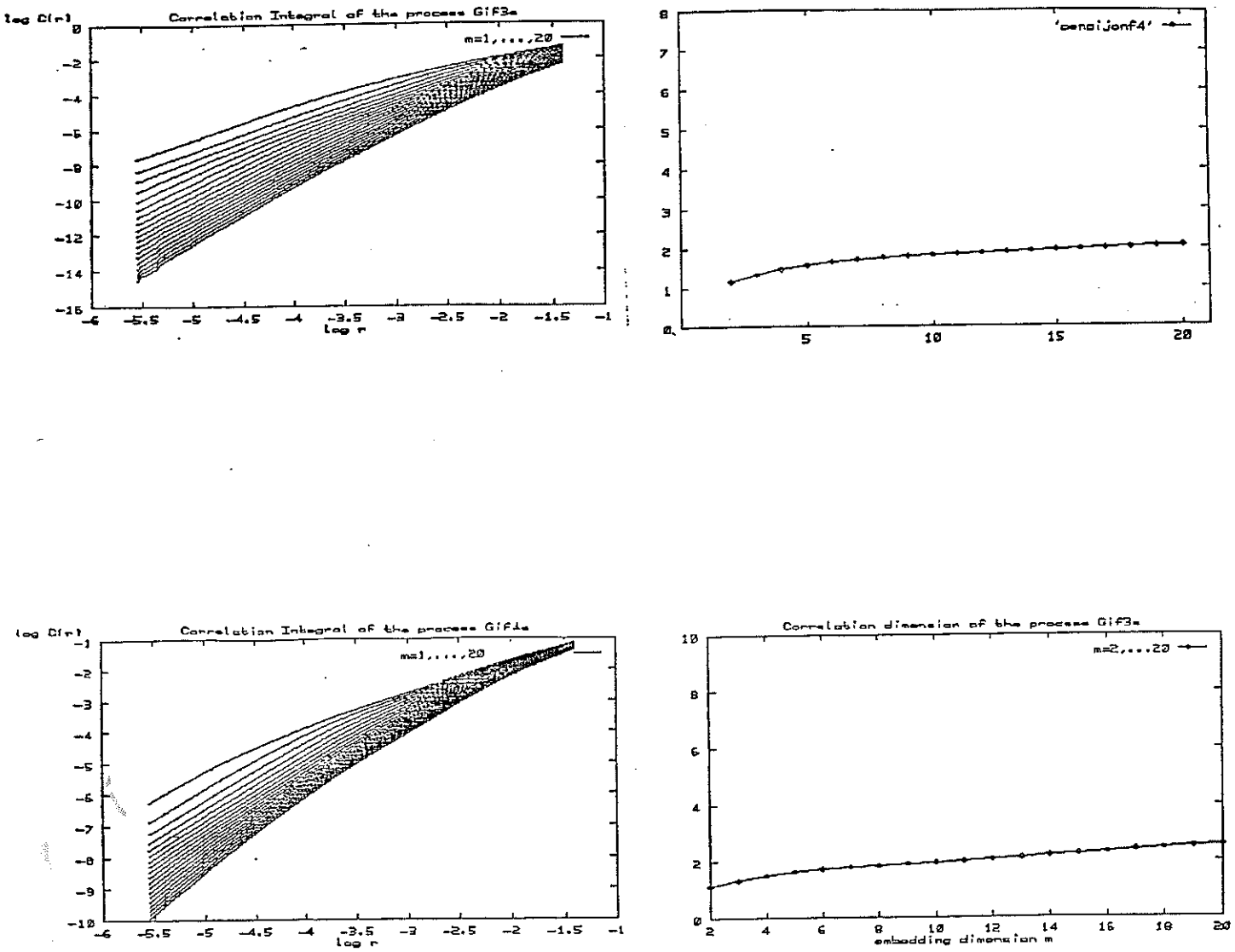
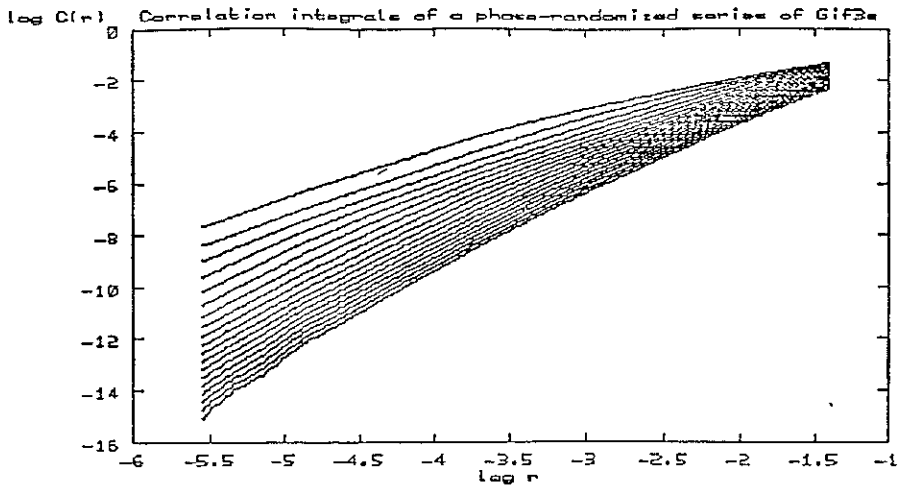


Figure C.43: *Correlation integrals and correlation dimension of the filtered SWH-series $Gif3_S$ and $Gif4_S$.* The correlation integrals and the correlation dimension estimate $\nu(m)$ from m=2 up to m=20 are shown for the series $Gif3_S$ (top) and $Gif4_S$ (bottom). Notice the apparent convergence of the values $\nu(m)$ to $\nu \simeq 2$ for the series $Gif4_S$ (the same effect also appears in the analysis of $Gif3_S$, but more weakly). This is seemingly consistent with the SAH. Further testing will show, however, that this is not the case. Due to the inherent noise introduced by the filtering process it was not necessary to add a small noise in order to analyze the filtered data. (CPU times=67.74 and 114.2 min.)



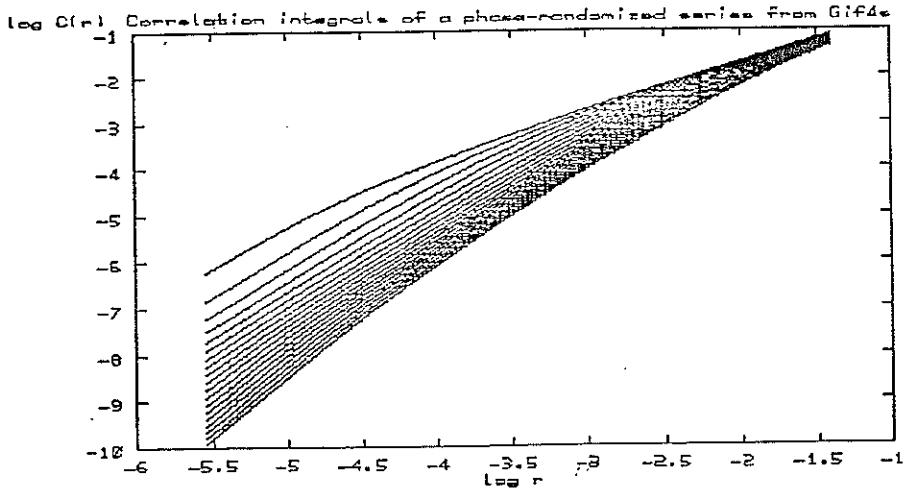


Figure C.44: *Correlation analysis of phase-randomized filtered SWH-series.* This figure shows the correlation integrals of $Gif3_S$ and $Gif4_S$ series with their Fourier phases randomized. This test is explained in appendix A. where its role in detecting spurious finite-dimensional estimates is pointed out. The fact that the correlation integral behaviour of the surrogate series do not differ from that of the original series supports the rejection of the SAH. (CPU times=65.9 and 115.95 min.)

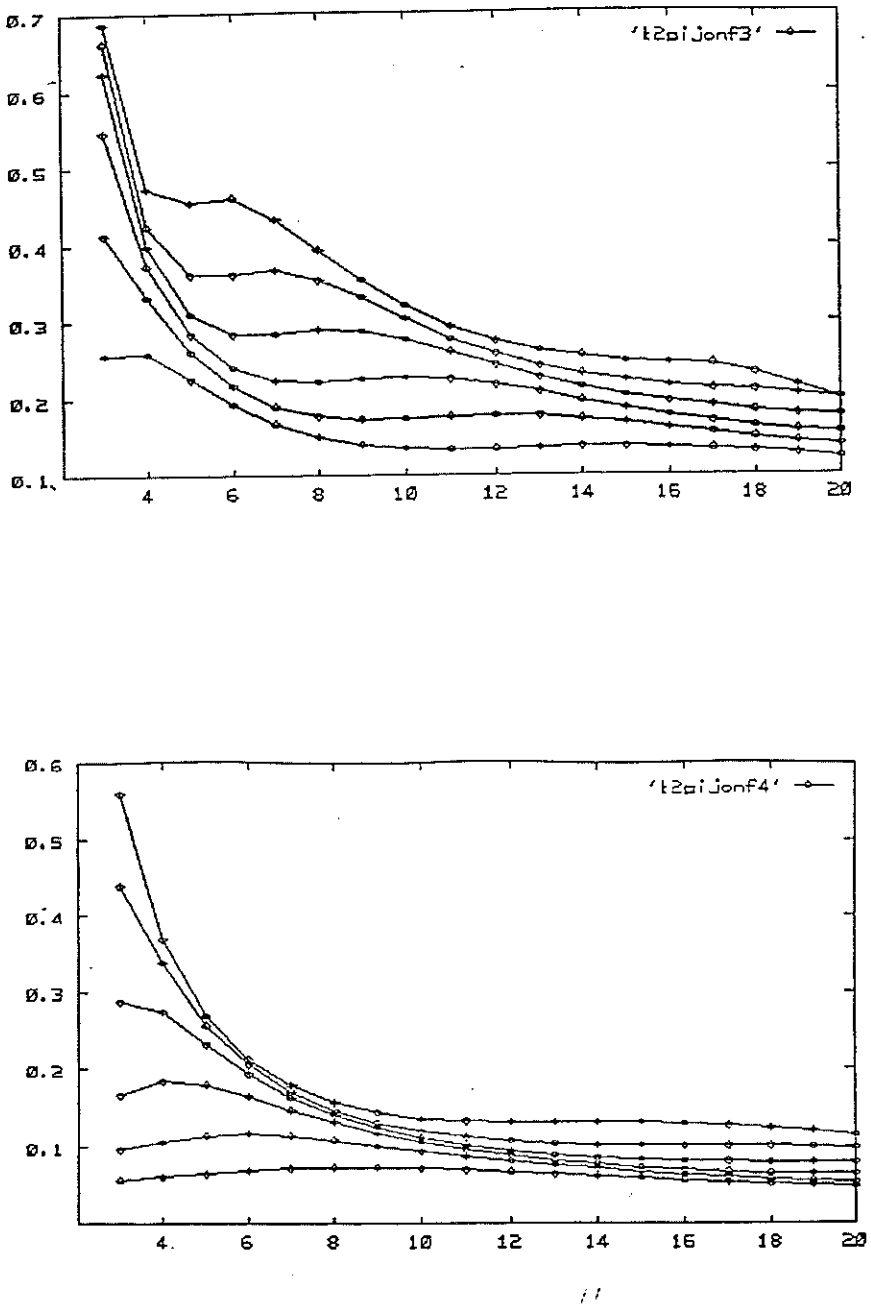
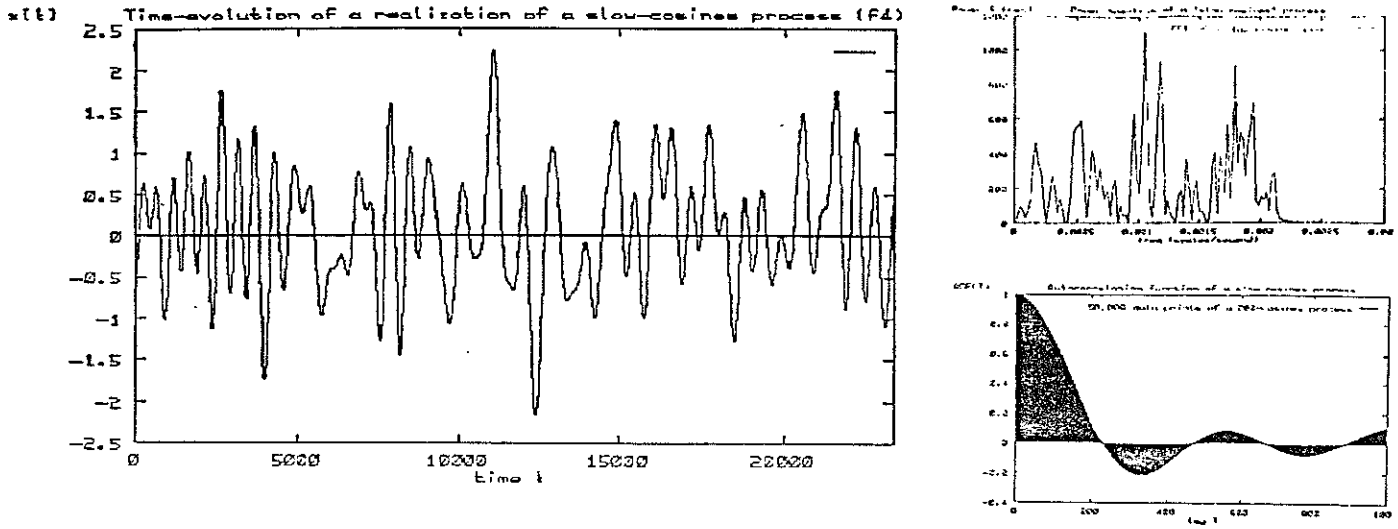


Figure C.45: *Entropy Curves of the filtered series $Gif3_S$ and $Gif4_S$.* The curves $K_2(\varepsilon, m)$ are plotted for six values of ε up to $m = 20$ (going further up is not allowed by the scarcity of data). For small ε's values (upper curves) there is a significant decrease in the randomness of this process with respect to that present in the unfiltered process Gi_S (see figure C.38). This seems to be in accordance with the SAH, mistakenly as further testing shows (see chapters 3 and 4).

C.14 Slow–Cosines Processes. Correlation and Entropy Analysis.



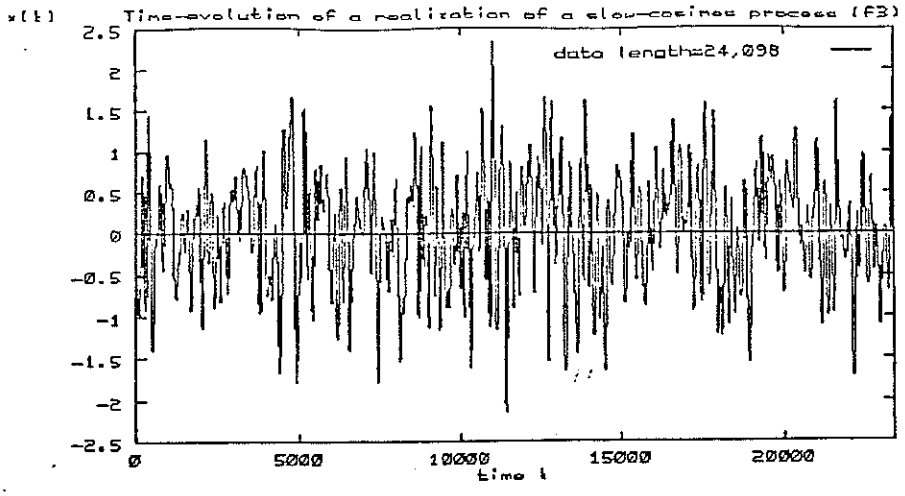


Figure C.46: *Slow-cosines series.* Realization of a slow-cosines stochastic process (see section **3.2.2** with 200 phase-random cosines, and crude power spectrum (FFT) of the process. The frequencies of the cosines were chosen equidistributed within the frequency interval of the series $Gif4_S$ (i.e. $(0, 0, 00215)$), so we denote this process as f4–cosines process. The autocorrelation function (ACF) is also shown. A slow–cosines process with frequencies equidistributed on the range of the $Gif3_S$ series, which we denote by f3-cosines process, is shown below.

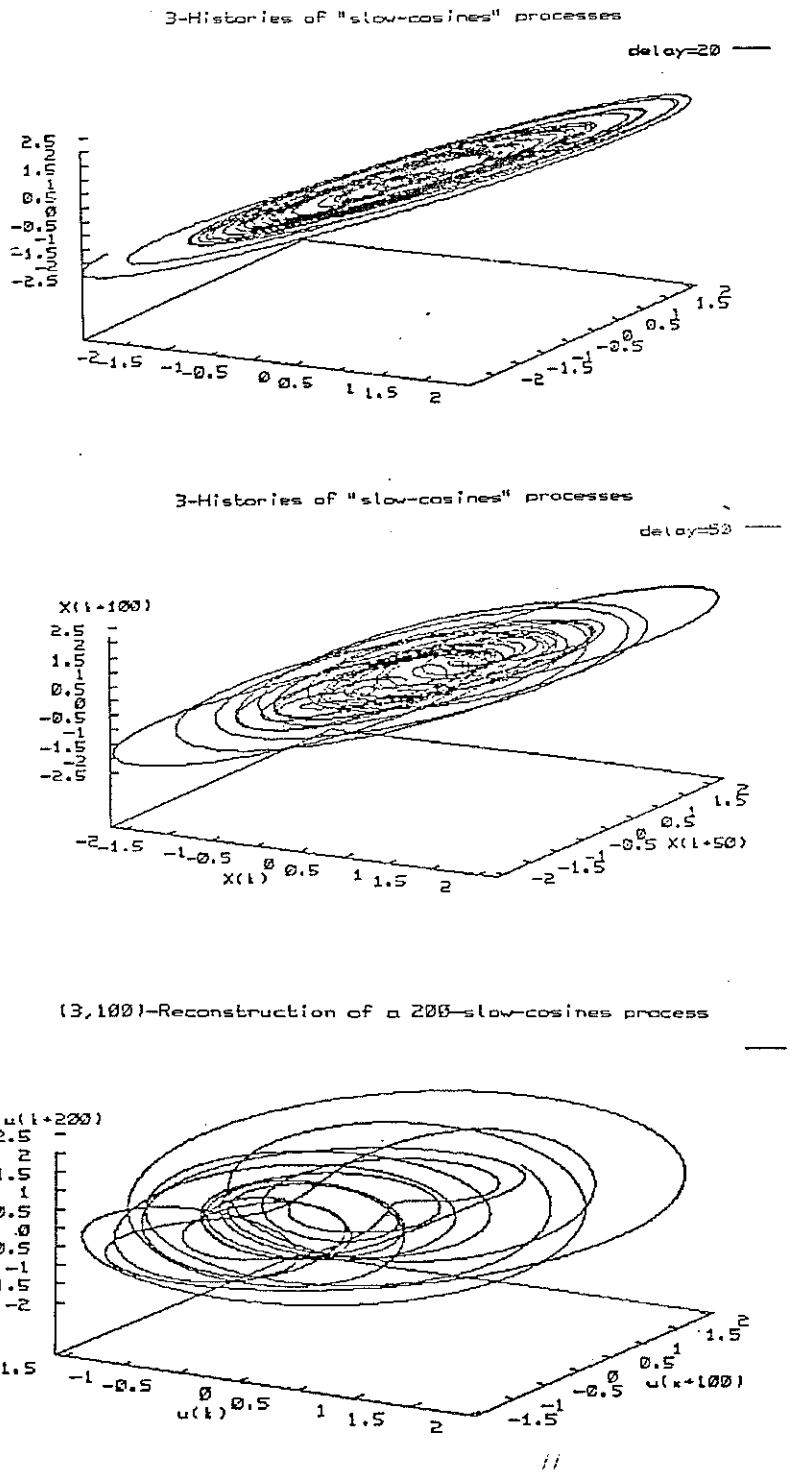
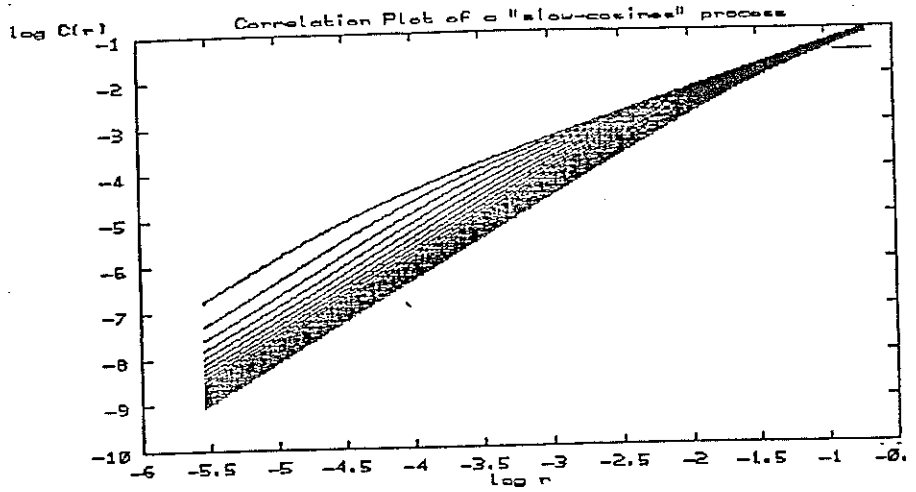


Figure C.47: *Delay 3-dimensional reconstructions of a realization of a slow-cosines stochastic process.* We considered delay steps $\tau = 20, 50$ in order to avoid the stretching of the data set along the line $x = y = z$. An optimal delay plot (according to the criterion in section 1.3 is obtained by using $\tau = 100$. as can be seen from the autocorrelation function of the series (figure C.46).



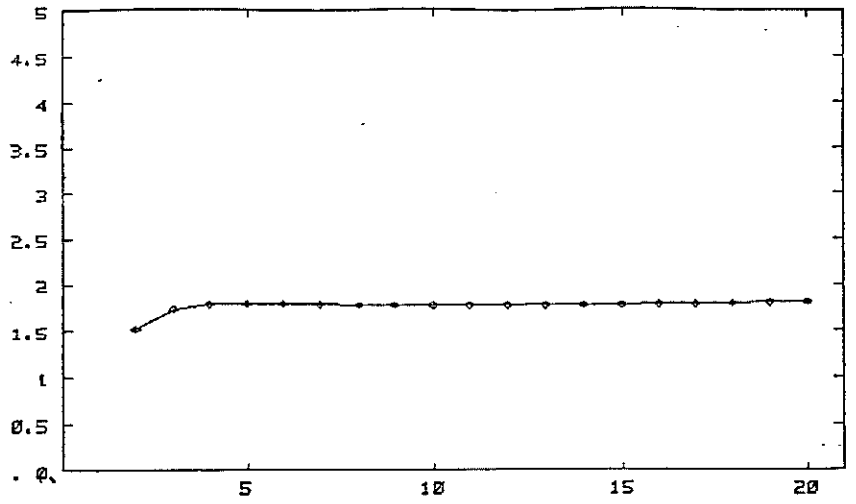
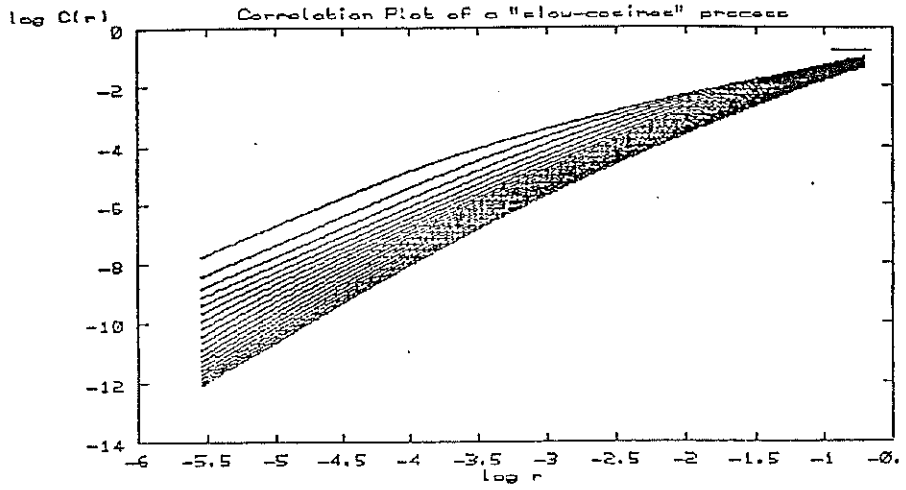


Figure C.48: *Correlation integrals of a f4-cosines process.* The figure shows the correlation integrals and the correlation dimension estimates ($m = 2, \ldots, 20$) for a realization of length 24,098 of the f4-cosines process (see C.46). In spite of the white randomness of the process, the correlation integrals provide a well-defined very low estimate for the correlation dimension $\nu \simeq 1.8$. This seems to be a feature of finite-length dimension estimates for low-frequency content processes. (CPU time=128.9 min)



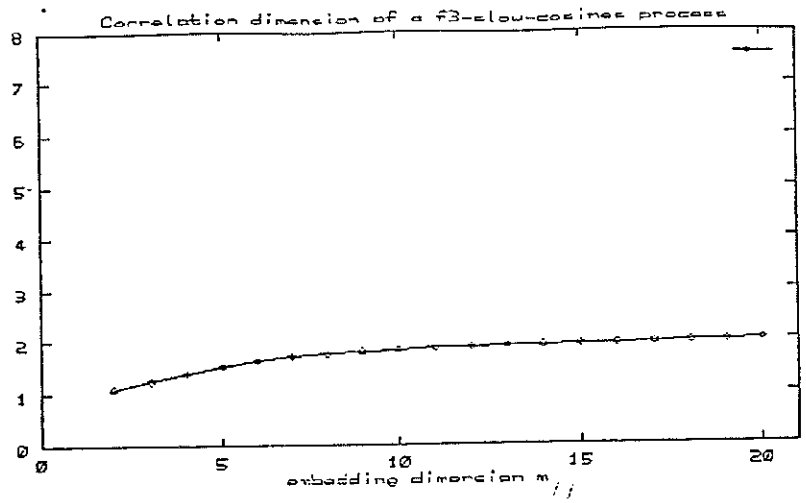
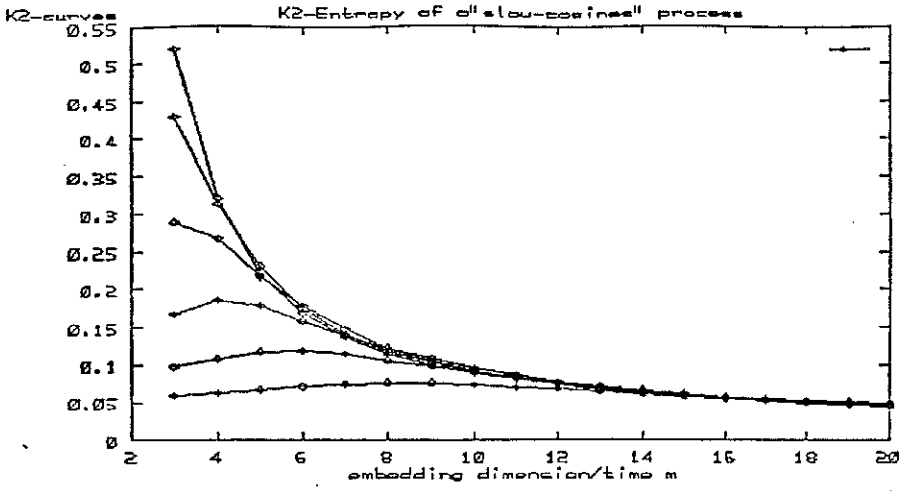


Figure C.49: *Correlation integrals of a f3-cosines process.* This figure shows the same analysis as in the previous figure (C.43) for a realization of length 24,098 of a slow-cosines stochastic process with frequencies equidistributed in the interval (0, 0.0065) (as the series $Gif3_S$). The results are very similar to those of C.43. (CPU time=118.252)



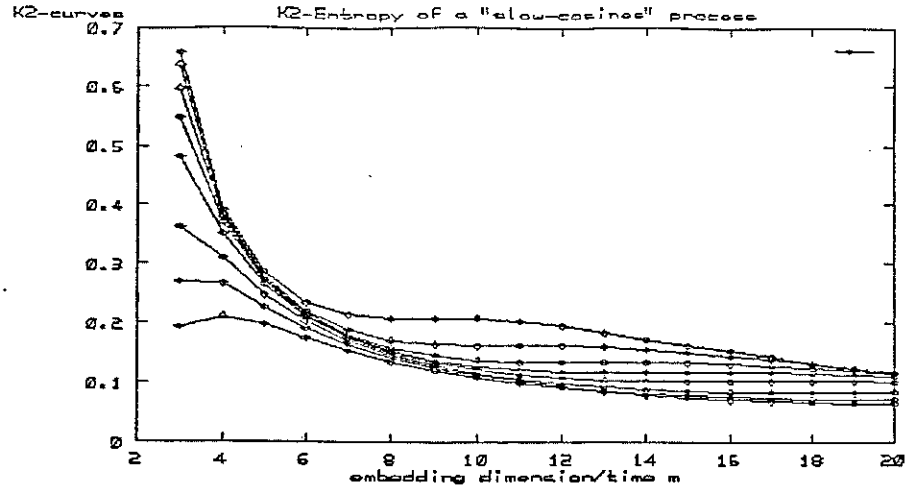


Figure C.50: *Entropy curves of slow-cosines processes.* Entropy plots of the f4–cosines process (above) given in C.46, and of the f3–cosines process (below) whose correlation integrals are given in figure C.49. Notice the fast convergence to an estimate value $\simeq 0.05$ in the case of the f4–cosines. This example provide a way of explaining the finite estimates obtained for the series $Gif4_S$, thus denying the possibility of accepting the SAH.

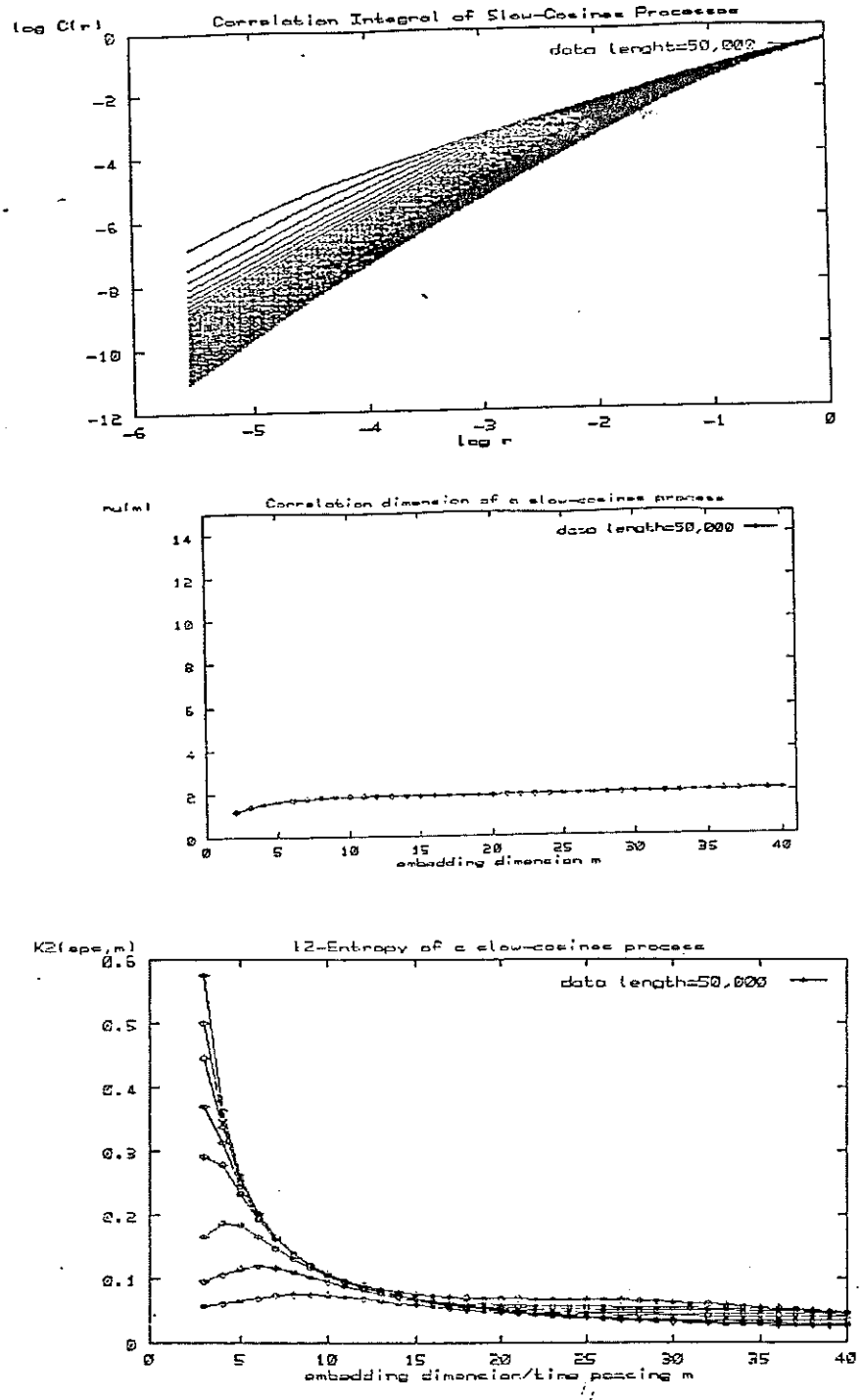


Figure C.51: *Correlation integrals and entropy plots of a f4-cosines process with* 50.000 *data points.* We duplicate the length of a series generated from a f4–cosines process like that in C.46, and compute the correlation integrals in order to exhibit the severity of the finite estimates of the correlation dimension of slow-cosines processes. The entropy plot shows the same limitation for the entropy estimate, though a very timid increase of the estimates $K_2(\varepsilon)$ is appreciated for small values of ε.

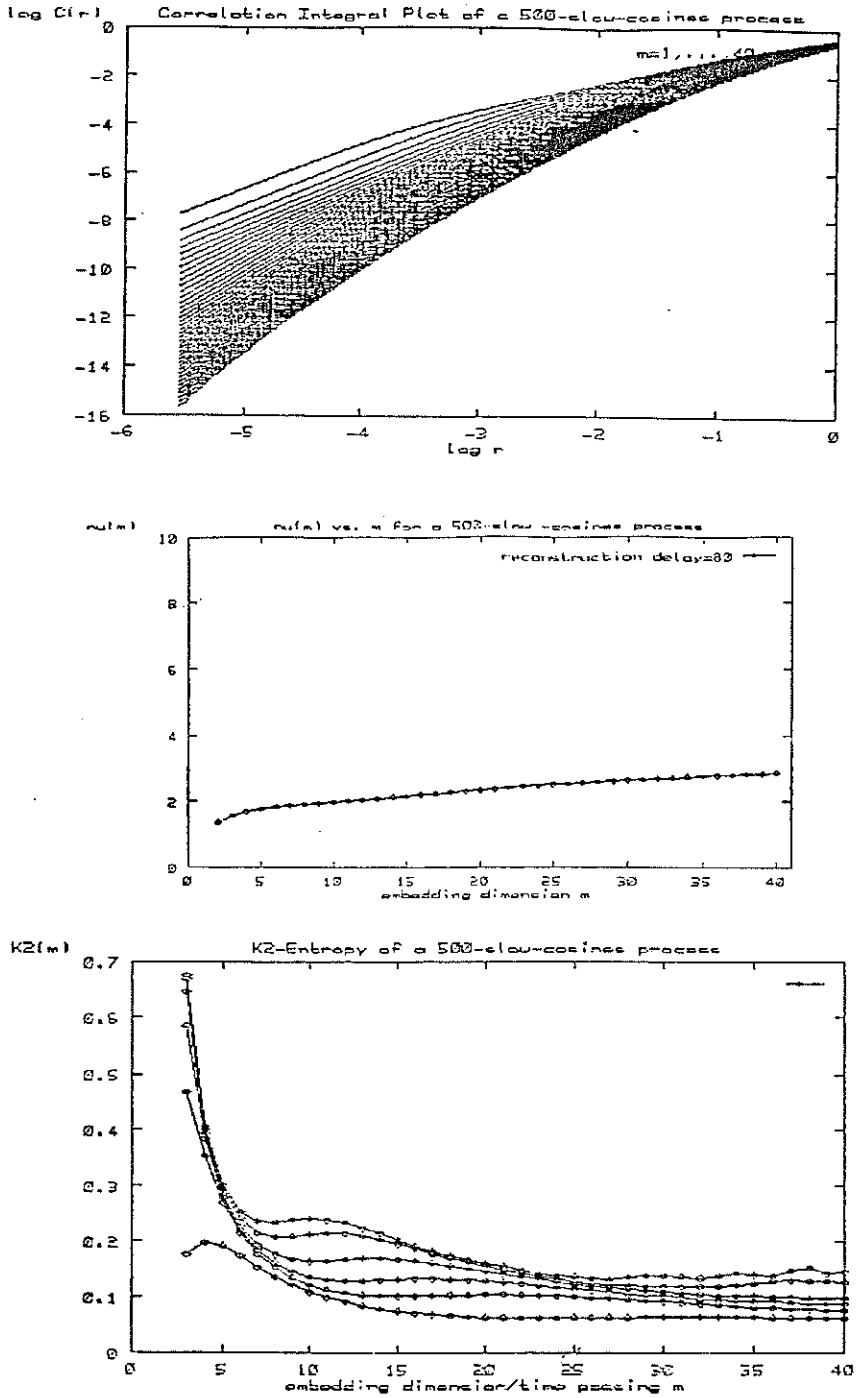


Figure C.52: *Analysis of a f3-cosines process with 500 cosines.* This figure shows the correlation and entropy plots of a slow-cosines process like that in C.49, but with 500 frequencies equidistributed in the frequency range of the series $Gif3_S$. We generated 50,000 points and went up to dimension 40. The computation was performed for several cut-off delays, that gave the same results. Both the correlation plots and the entropy curves show again the severity of the finiteness of the estimates for these kinds of processes.

C.15 Sensitivity and Stability Features of the Proposed Algorithm to Compute Liapunov Exponents with respect to the Input Parameters .

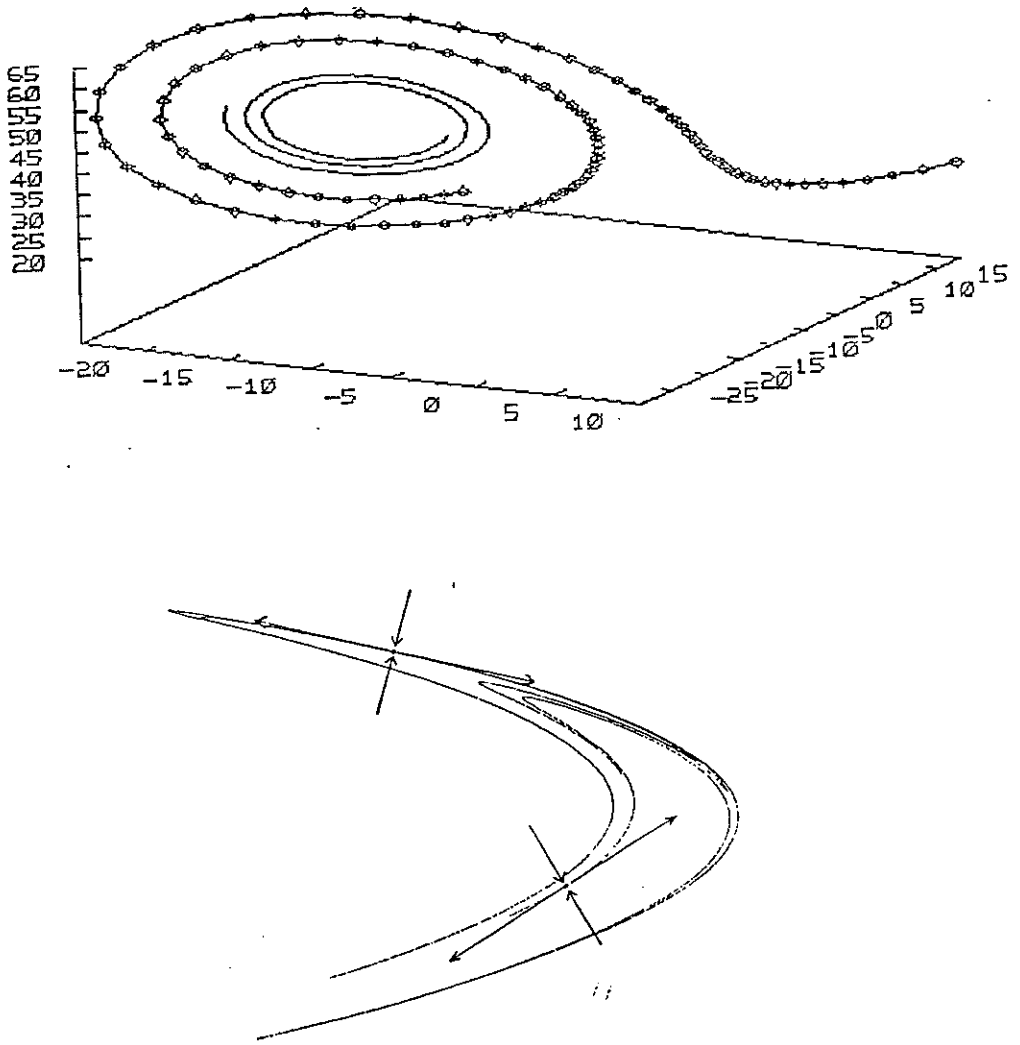


Figure C.53: *A criterion for fixing the radius ρ of balls where the linear fits are performed when estimating the Liapunov exponents:* the orbits of the flow appear as straight line segments when observed in that scale. The top figure illustrates this criterion for the Lorenz attractor. The bottom figure shows how the unstable manifold of the Henon attractor (the parabolic–shaped black lines) can replace the orbits in the case of discrete–time flows.

ρ	2	1	0.9	0.8
λ_1	1.4891	1.5990	1.6314	1.7191
λ_2	0.0263	-0.0053	-0.1293	0.2622
λ_3	-26.351	-14.880	-13.796	-11.310
$r(\rho,2)$	0.9994	0.9993	0.9990	0.9990

Table C.1: *Dependence of the negative Liapunov exponent of the Lorenz attractor on the radius* ρ. The table shows the computation of the Liapunov exponents (rows 2nd to 4th) computed over balls of different radii (1st row) for an orbit of 5,000 data points of the Lorenz attractor generated by a fourth order Runge–Kutta scheme with integration step $\Delta\tau = 0.01$ and with sampling time $\Delta t = 0.03$ (see C.2). The number of iterations $N_{\rm it}$ was fixed to 1,000. Whilst the positive and null exponents are rather stable, the negative exponent shows an extremely sensitive dependence on the radius, which illustrates the difficulty of computation of negative exponents in a situation of lack of data. An increase of the radius often improves the estimates, even though it reinforces the weight of the non–linear second–order terms in the power expansion of the flow. (CPU time=0.46 min.)

$N_{\rm orb}$	5,000	10,000	50,000
λ_1	1.5990	1.5304	1.5133
λ_2	-0.0053	-0.0422	-0.0249
λ_3	-14.880	-20.963	-24.934
$r(\rho,2)$	0.9993	0.9998	0.9998

$N_{\rm orb}$	5,000	10,000	50,000
λ_1	1.7191	1.5041	1.5172
λ_2	0.2622	0.0209	-0.0273
λ_3	-11.310	-18.589	-23.678
$r(\rho,2)$	0.9990	0.9998	0.9998

Table C.2: *Dependence of the Liapunov exponents of the Lorenz attractor on the radius* ρ *and the number of data points* $N_{\rm orb}$. The parameters of the algorithm have been predetermined with the same values as in table C.1 except for the radius; in the top table the radius is fixed at the value of 1, and in the bottom table at 0.8; the number of data points ($N_{\rm orb}$) is given by the first row of each table. One can verify in this way that, for increasing values of $N_{\rm orb}$, the negative Liapunov exponent decreases, in both cases, down to the true value (-22.5). Further increases of $N_{\rm it}$ do not result in estimates closer to -22.5; on the contrary, they result in a further lessening of the value of the exponent. The estimate obtained in this way actually converges to the limit value $\lambda(\rho,2)$ $\rho = 1,\ 0.8$) of the negative Liapunov exponent when computed from linear fits obtained over balls of radius ρ; this value differs from the one obtained when using the maps tangent to the flow. This fact illustrates the dependence of Liapunov exponents on the radius of the balls. (CPU time=1.80 min.)

ρ	1	0.9	0.8	0.7	0.6
λ_1	1.5133	1.4658	1.5172	1.5566	1.5290
λ_2	-0.0249	0.0244	-0.0274	-0.0155	0.0206
λ_3	-24.934	-24.660	-23.678	-22.555	-21.844
$r(\rho, 2)$	0.9998	0.9999	0.9999	0.9999	0.9999

ρ	1	0.9	0.8	0.7	0.6
λ_1	1.4933	1.4890	1.4994	1.5126	1.5061
λ_2	-0.0069	0.0212	-0.0275	-0.0041	0.0065
λ_3	-25.205	-25.205	-23.371	-22.173	-21.384
$r(\rho, 2)$	0.9999	0.9999	0.9999	0.9999	0.9999

Table C.3: *Stability of the estimates of Liapunov exponents of the Lorenz attractor with respect to the number $N_{\rm it}$ of iterations.* The values of the parameters in both tables are: length of the orbit, $N_{\rm orb}$ =50,000, sampling time, Δt=0.03, starting point 1,000. The Liapunov exponents are computed in the top table for a number of iterates $N_{\rm it}$=10,000, and in the bottom table for $N_{\rm it}$=5,000, with slight differences in both cases. (CPU time=11.55 min.)

ρ	1	0.9	0.8	0.7	0.6
λ_1	1.4994	1.4788	1.5198	1.5182	1.5145
λ_2	-0.0338	0.0165	-0.0422	-0.0072	0.0003
λ_3	-25.488	-25.446	-24.818	-23.630	-22.726
$r(\rho, 2)$	0.9992	0.9995	0.9995	0.9997	0.9998

Table C.4: *Stability of the Liapunov exponents in the Lorenz attractor with respect to sampling time.* Compare this with table C.1, corresponding to identical values for all the code parameters, except for the sampling time Δt, that was equal to 0.03 in that table, and takes a value of 0.05 in this table.

ρ	1	0.8	0.7
λ_1	1.5775	1.5119	1.5419
λ_2	-0.0066	0.0019	-0.0240
λ_3	-25.127	-23.477	-22.544
$r(\rho, 2)$	0.9999	0.9999	0.9999

Table C.5: *Stability of the estimates of the Lorenz Liapunov exponents with respect to the starting point.* In this table we pick the 20,000th point of the orbit as the starting point. Compare this with the top table in C.3, computed with the 10,000th point as starting point. All other parameters are the same in both tables.

C.16 Analysis of the Liapunov Spectrum Computation for the SWH–Series.

ρ	0.5	0.4	0.3	0.2	0.1	0.05
λ_1	-0.027	-0.027	-0.027	-0.027	-0.046	-0.013
$r(\rho,2)$	0.515	0.423	0.312	0.187	0.066	0.029

Table C.6: *Liapunov exponents for the univariate series Gi_S.* The Liapunov exponents are computed from linear fits performed over balls of various radii, which can be read on the first row of the table. The second row gives the values of the Liapunov exponent. The third row gives values for the parameter $r(\rho,2)$, evaluating the accuracy of the linear fits. According to this parameter, the linear fits are poor and they worsen when the radii decrease. The Liapunov exponents do not measure orbital divergence. Thus, their negative values cannot be interpreted as a sign of a deterministic dynamics. (CPU time=73.31 min.)

ρ	0.5	0.4	0.3	0.2	0.1	0.05
λ_1	-0.021	-0.020	-0.020	-0.020	-0.048	-0.034
$r(\rho,2)$	0.527	0.431	0.317	0.186	0.0655	0.028

Table C.7: *Stability of the Liapunov exponent of the series Gi_S with respect to the starting point.* The computation of the Liapunov exponent is made for a starting point different from that used in table C.6. There, the 4,450-th point of the orbit was selected as the starting point. Here the starting point is the 20,000-th. The coincidence of the results provides evidence of strong ergodicity of the observed stochastic process with respect to the computed parameters. (CPU time= 85.5 min.)

ρ	0.5	0.4	0.3	0.2	0.1	0.05
λ_1	-0.018	-0.019	-0.020	-0.022	-0.040	
$r(\rho,2)$	0.584	0.494	0.379	0.230	0.085	

Table C.8: *Stability of the Liapunov exponent of the SWH-series Gi_S with respect to the number of iterations.* The Liapunov exponent of the series Gi_S is computed for the same starting point as in table C.7, with an increase of the number N_{it} of linear fits from a value of 1,000 (in tables C.6 and C.7) up to a value of 2,000. The values of the Liapunov exponent and those of the parameter $r(\rho,2)$ are stabilized with respect to such an increase. In table C.13 (compare with the first column in the table above) one can see that a much lower number of iterations (N_{it} =100) does not guarantee the required stability. (CPU time=177.63 min.)

ρ	0.5	0.4	0.3	0.2	0.1	0.05
λ_1	-0.041	-0.038	-0.037	-0.032	-0.001	-0.119
λ_2	-0.418	-0.403	-0.389	-0.365	-0.315	-0.235
$r(\rho,2)$	0.569	0.492	0.399	0.283	0.149	0.102

Table C.9: *Computation of the Liapunov spectrum of the 2-histories of the Gi_S series.* The largest and second Liapunov exponents can be read from rows two and three respectively. The point made in the legend of table C.6 applies to this case. (CPU time=64.04 min.)

ρ	0.5	0.4	0.3	0.2	0.1	0.05
λ_1	-0.024	-0.021	-0.011	0.000	0.059	0.148
λ_2	-0.180	-0.167	-0.148	-0.129	-0.068	0.007
λ_3	-0.488	-0.458	-0.457	-0.409	-0.324	0.191
$r(\rho, 2)$	0.595	0.529	0.436	0.330	0.234	0.221

Table C.10: *Liapunov exponents of the 3-histories of the series Gi_S.* The first, second, and third Liapunov exponents can be read from rows one, two, and three respectively. (CPU time=64.68 min.)

ρ	0.5	0.4	0.3	0.2	0.1	0.05
λ_1	-0.025	-0.023	-0.017	-0.001	-0.055	0.170
λ_2	-0.180	-0.172	-0.153	-0.121	-0.075	0.007
λ_3	-0.482	-0.470	-0.457	-0.427	-0.293	-0.167
$r(\rho, 2)$	0.598	0.527	0.434	0.325	0.210	0.230

Table C.11: *Liapunov spectrum of the 3-histories of the series Gi_S from a different starting point.* In tables C.8, C.9, and C.10 the point 4,450 was selected as starting point. Here, the starting point is the 20,000-th, without a significant change in the results (compare this with table C.10).

Dim	1	2	3	4	5	6	7	8	9	10
λ_1	-0.026	-0.041	-0.024	-0.107	-0.005	-0.001	0.003	0.002	0.007	0.006
λ_2		-0.418	-0.180	-0.899	-0.051	-0.039	-0.027	-0.022	-0.020	-0.016
λ_3			-0.488	-0.288	-0.170	-0.139	-0.114	-0.094	-0.076	-0.060
λ_4				-0.409	-0.204	-0.162	-0.139	-0.108	-0.090	-0.072
λ_5					-0.263	-0.207	-0.162	-0.122	-0.104	-0.083
λ_6						-0.327	-0.198	-0.151	-0.114	-0.091
λ_7							-0.277	-0.178	-0.129	-0.101
λ_8								-0.250	-0.151	-0.111
λ_9									-0.199	-0.132
λ_{10}										-0.206
$r(\rho, 2)$	0.51	0.56	0.59	0.61	0.63	0.64	0.66	0.67	0.68	0.69

Table C.12: *Liapunov spectrum of the m-histories of the series Gis for embedding dimensions* $m = 1, \ldots, 10$. The exponents can be read from the columns 2nd to 11th of the table. The first column gives the value of m. The last column gives the value of the estimate of the correctness of the linear fits. In accordance to the anomalous negative values of the Liapunov exponents, the linea fits are poor. The point 4,450-th was selected as a starting point, the radius of the balls where the linear mappings are fitted was 0.5, and the number of iterations (linear fits) is 1,000. (CPU time=325.64 min.)

Dim	1	3	6	9	12	15
λ_{max}	-0.021	-0.011	0.000	0.000	0.006	0.004
λ_{min}		-0.122	-0.119	-0.110	-0.095	-0.087
$r(\rho,2)$	0.060	0.179	0.258	0.320	0.362	0.354

Dim	1	3	6	9	12	15
λ_{max}	-0.046	-0.018	0.002	0.018	0.005	0.004
λ_{min}		-0.423	-0.284	-0.170	-0.258	-0.100
$r(\rho,2)$	0.393	0.495	0.597	0.672	0.700	0.700

Table C.13: *Dependence of the accuracy of the linear fits on the sampling time for the series* Gi_S. The sampling time selected in the top table was Δt =10. The number of linear fits was lessened to 100 in order to avoid the gaps of missing data. The second and third rows show, respectively, the maximum and Liapunov exponents of the m-histories for the values of m corresponding to the first row. The last row gives the values of the parameter $r(\rho,2)$. If we compare these values with the corresponding values in the bottom table above, where the selected sampling time was Δt =1, we learn that the accuracy of the linear fits diminishes when the sampling time increases. We can observe in these tables that the values of $r(\rho,2)$ do not stabilize for the given number of iterations (compare this with the first column of tables C.6 and C.8). (Value of the radius ρ=0.5).

ρ	0.25	0.015	0.05	0.025	0.0075	0.0025
λ_1	0.44710	0.43830	0.43446	0.42960	0.43300	0.43117
λ_2	-1.6321	-1.6438	-1.5490	-1.6191	-1.6286	-1.6286
$r(\rho,2)$	0.95984	0.98646	0.99601	0.99551	0.99976	0.99976

Table C.14: *Computation of the parameter $r(\rho,2)$ for the Henon attractor.* When the radii of the balls, where the linear fits are performed, decrease, the parameter approaches to one, thus showing the decay of the weight of the second order terms in the power expension of the flow. (CPU time= 21.75)

ρ	2	0.6	0.4
λ_1	1.414797	1.528962	1.4931339
λ_2	-0.0090	0.020596	0.008736
λ_3	-30.141	-21.844	-19.108
$r(\rho,2)$	0.999287	0.999941	0.999978

Table C.15: *Computation of the parameter $r(\rho,2)$ for the Lorenz attractor.* The computation is performed over an orbit of 50,000 points generated by fourth-order Runge–Kutta scheme with integration step $\Delta\tau$=0.01. The orbit has a length ($N_{\rm orb}$) of 50,000 points. The sampling time is Δt=0.03, and the number of iterations ($N_{\rm it}$) is 1,000. Observe that although the smallest radius ρ=0.4, which allows the best linear fits, gives the best estimates for the first and second Liapunov exponents, it gives a worse estimate for the third Liapunov exponent than the radius ρ =0.6. This is due to the scarcity of data points in balls of radius 0.4 (10 data points in the average, whilst the averaged number of data points in balls of radius 0.6 is 29).

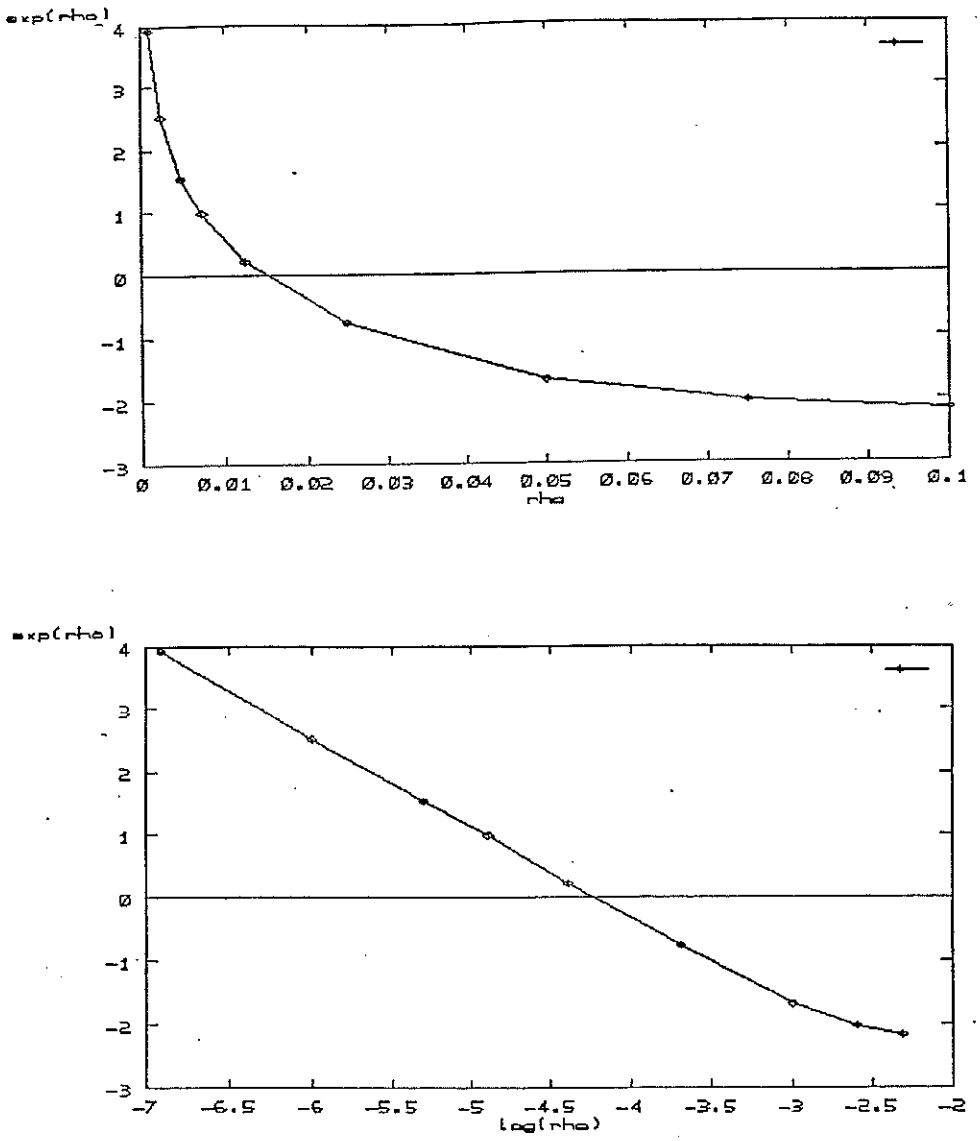


Figure C.54: *Liapunov exponent of an i.i.d. sequence of random variables uniformly distributed over [0,1].* A plot of the Liapunov exponents vs. radii shows an increase of the Liapunov exponents with the decrease of the radii. The negative values of the exponent in the range [0.02.0.1] of the radii shows that these Liapunov exponents do not measure divergence between nearby orbits. This is consistent with the low values of the parameter $r(\rho.2)$ shown in tables C.16 and C.17 below. The plot of the Liapunov exponent vs. $\log \rho$ of the same stochastic process evidences the existence of a power–law dependence between the radii of the balls and the expansion ratio of the averaged linear fit over such balls. The length of the orbit generated for this computation was 25,000, and the number of iterations was 1,000.

ρ	0.1	0.05	0.025	0.005	0.0025	0.0011
λ_1	-2.1599	-1.6747	0.223647	1.533827	2.5210	3.92877
$r(\rho,2)$	0.02200	0.01146	0.003570	0.00468	0.00802	0.02073

Table C.16: *Estimates of the Parameter $r(\rho,2)$ for an i.i.d. sequence of random variables uniformly distributed over [0,1].* The smalll values of the parameter indicate poor linear fits for this process. (CPU time=4.64)

ρ	0.1	0.05	0.025	0.0125	0.0075	0.0025
λ_1	-0.4820	0.016819	0.9359	2.15465	3.01820	1.902
λ_2	-1.1867	-0.6529	-0.2481	-0.0309	0.2759	0.816
$r(\rho,2)$	0.004653	0.029745	0.041515	0.14091	0.3776	0.5273

Table C.17: *Liapunov exponents of the 2-histories of a sequence of i.i.d. random variables uniformly distributed over the interval [0,1].* The behaviour of the Liapunov exponents is similar to that of the univariate series, with values of the first and second exponents increasing with decreasing radii. The parameter $r(\rho,2)$ shows the linear fits to be very poor. The high value of this parameter for the smallest radii is due to the scarcity of available data points for the fitting. For a radius of 0.005 the number of data points for which the linear fits could not be performed because of a lack of neighbours (i.e. less than two) was 531, from a total number of 1,000 iterations. The special device designed in our code (see 4.3.2) for such a case allowed us to compute the Liapunov exponent. (CPU time=4.6 min.)

ρ	0.25	0.2	0.15	0.1	0.05	0.01
λ_1	0.000	0.000	0.000	0.000	0.000	-0.002
$r(\rho,2)$	0.995	0.993	0.987	0.972	0.901	0.339

Table C.18: *Liapunov exponents of the dynamics of the series* $Gif4_S$. We use a sampling time Δt =1, which causes the flow to behave approximately as the identity mapping, given the low frequencies of the spectral components of this series. The computed Liapunov exponents are thus zero, except for the smallest radius, where the computation becomes more sensitive to the small variations of the flow. Notice that for this radius the value of the parameter $r(\rho,2)$ falls abruptly. (CPU time= 52.69)

ρ	0.25	0.2	0.15	0.1	0.05
λ_1	0.000	0.000	0.000	0.000	0.000
$r(\rho,2)$	0.997	0.995	0.992	0.982	0.935

Table C.19: *Liapunov exponents for a phase-randomized series* $Gif4_S$. The Liapunov exponents and fitting parameter $r(\rho,2)$ obtained are very similar to those obtained for the $Gif4_S$ series over the same range of radii. This shows that the null values of the Liapunov exponents of the $Gif4_S$ series cannot be attributed to a deterministic behaviour.

Dim.	1	2	3
λ_1	-0.002	0.000	0.000
λ_2		-0.004	0.000
λ_3			-0.003
$r(\rho,2)$	0.083	0.703	0.976

Table C.20: *Liapunov exponents of the m-histories of the series $Gif4_S$.* The sampling time has been raised to 50 in order to counterbalance the small frequencies of the filtered components. This permits only 20 linear fits ($N_{\rm it}$=20). The radius ρ is given a value of 0.2. There appear non-null values for some Liapunov exponents. When the embedding dimension attains the value 2, which is larger than the value of the correlation dimension (see figure C.43), the value of $r(\rho,2)$ approaches one. Thus the series $Gif4_S$ emulates thus one more time a deterministic dynamics living in dimension two.

Dim.	1	2	3
λ_1	-0.0068	0.0003	0.0082
λ_2		-0.0012	0.0036
λ_3			-0.0067
$r(\rho,2)$	0.046	0.953	0.993

Table C.21: *The Liapunov spectrum for m-histories of an f_4-cosines process with sampling time Δt =100.* The qualitative behaviour of this process is quite similar to that of the series $Gif4_S$.

ρ	0.5	0.4	0.3	0.2	0.1	0.05
λ_1	0.034	0.051	0.079	0.129	0.239	0.313
$r(\rho, 2)$	0.876	0.851	0.816	0.761	0.681	0.641

Table C.22: *Estimate of the rate of orbital norm-divergence for the 3-histories of the series Gis.* We perform linear fits over the norms of the vectors in the tangent as explained in section 4.4.4c). The Liapunov exponent (second row) computed over balls with different radii (first row) now becomes positive, and the rightness of the linear fits improves with respect to those provided for the Eckmann-Ruelle algorithm. (CPU time= 46.14 min.)

Dim.	1	3	5	7	9	10
λ_1	0.03706	0.03399	0.02636	0.02147	0.01827	0.01704
$r(\rho, 2)$	0.6605	0.8755	0.9352	0.9589	0.9711	0.9750

Table C.23: *Estimate of the orbital norm-divergence of m-histories of the Gis series.* The embedding dimension can be read in the first row of the table. The linear fits become accurate for higher dimensions. The exponential rate of divergence decays slowly for increasing dimension. The radius taken was 0.5 for all dimensions. The number of linear fits performed was 1,000. (CPU time=146.02 min.)

Dim.	1	3	6	9	12	15
λ_1	0.081	0.058	0.033	0.023	0.015	0.011
$r(\rho, 2)$	0.610	0.858	0.945	0.964	0.976	0.985

Dim.	1	3	6	9	12	10
λ_1	0.025	0.031	0.031	0.029	0.025	0.020
$r(\rho, 2)$	0.464	0.562	0.632	0.677	0.710	0.741

Table C.24: *Dependence of the orbital norm-divergence of the series Gi_S on the number of iterations and the sampling time.* Both tables are computed for balls of a fixed radius equal to 0.5, the first one for 100 iterations and sampling time 1, and the second one for 100 iterations and sampling time 10. Compare both tables to test the dependence on the sampling time, and compare with table C.22 to test the influence of the number of iterations. The rate of divergence seems more more stabilized for sampling time 10, but the linear fits worsen in this case, though still are better that those from the uniform white noise (see table C.22).

ρ	0.25	0.2	0.15	0.1	0.05	0.025
λ_1	0.3924	0.3963	0.4016	0.4157	0.4029	0.3687
$r(\rho,2)$	0.90029	0.90734	0.9106	0.9146	0.91679	0.9306

ρ	2	1	0.9	0.8	0.7	0.6
λ_1	1.675820	1.664322	1.671073	1.681374	1.671402	1.661861
$r(\rho,2)$	0.985877	0.985882	0.985882	0.985911	0.986072	0.986016

Table C.25: *Estimates of the orbital norm–divergence of the Henon and Lorenz systems.* The algorithm used was the same as in tables C.22 to C.24 above. Observe that whilst the values obtained for the Henon attractor are slightly smaller than the maximum Liapunov exponent, in the case of the Lorenz attractor they are slightly larger. In both cases the linear fits are accurate. (CPU time=3.87 (Lorenz system) and 35.54 (Henon system)).

ρ	0.1	0.05	0.025	0.0125	0.005	0.0025
λ_1	1.595977	2.283365	2.972761	3.664710	4.577404	5.269270
$r(\rho,2)$	0.5208	0.5220	0.5230	0.5240	0.5262	0.5413

Table C.26: *Estimate of the orbital norm-divergence for an i.i.d. sequence of random variables uniformly distributed on the unit interval [0,1].* The estimates of the exponential rate of norm–divergence now become positive (compare this with the Liapunov exponent in table C.16), and the value of the parameter $r(\rho,2)$, although improved with respect to the quoted table, is still low. The plots show a remarkably exact potential law of dependence of the eigenvalue with respect to the radius, namely $\lambda = -1.0012407 \log \rho - 0.7094648$.

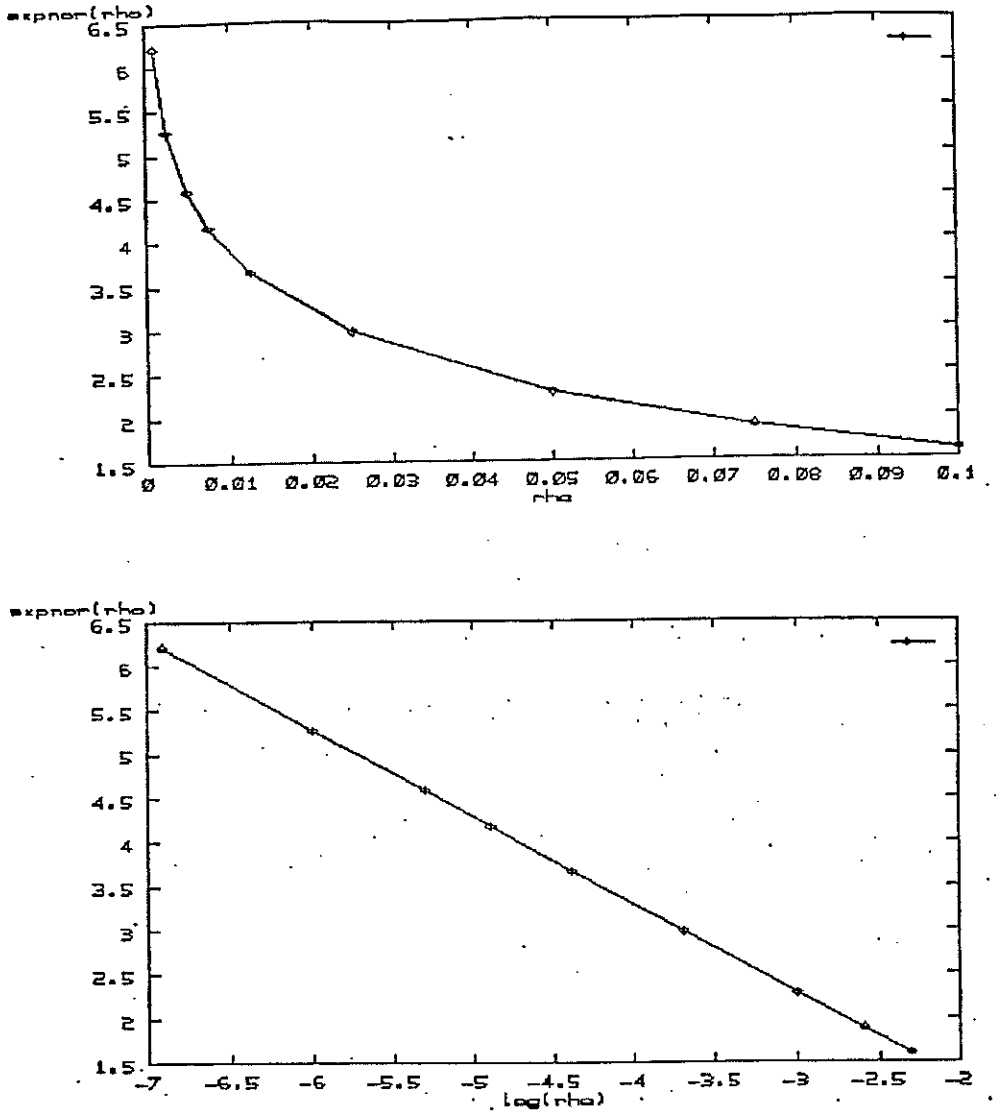


Figure C.55: *Power-law dependence with respect to the radius of the orbital norm-divergence of an i.i.d. sequence of random variables uniformly distributed over [0,1].* The bottom plot (linearly interpolated) is exactly filled by the function $\lambda = -1.0012407 \log r - 0.7094648$.

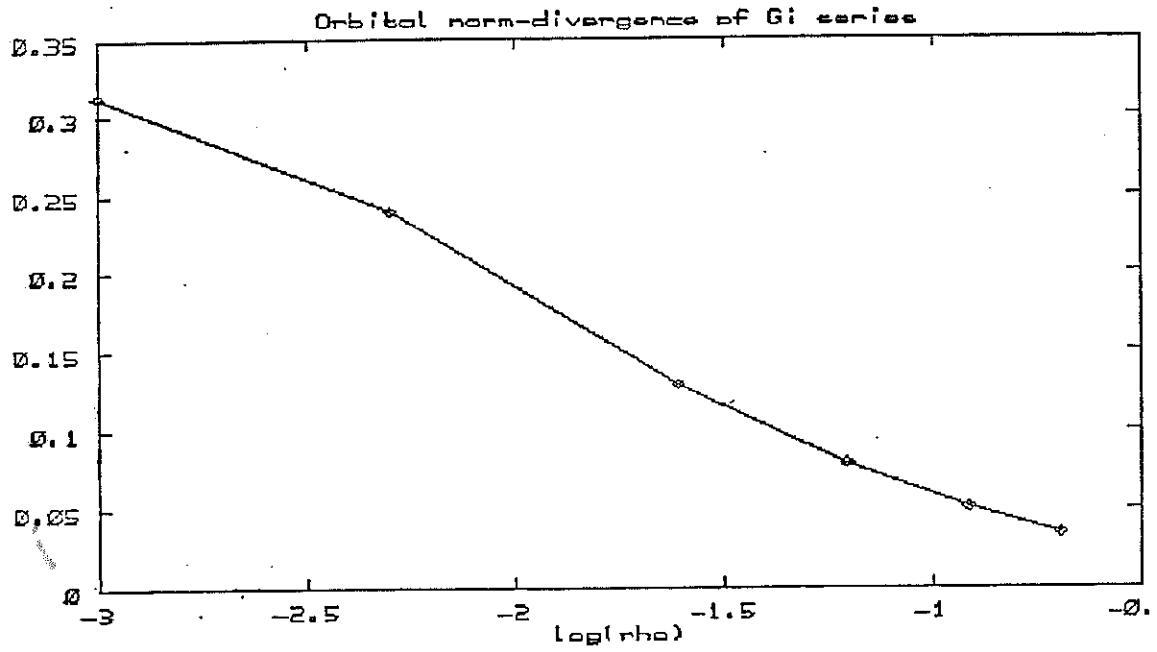


Figure C.56: *Power–law dependence with respect to the radius of the orbital–norm of the SWH–series Gi_S.* The plot shows a dependence for the Gi_S quite similar to that obtained in figure C.55 for a uniform white noise process.

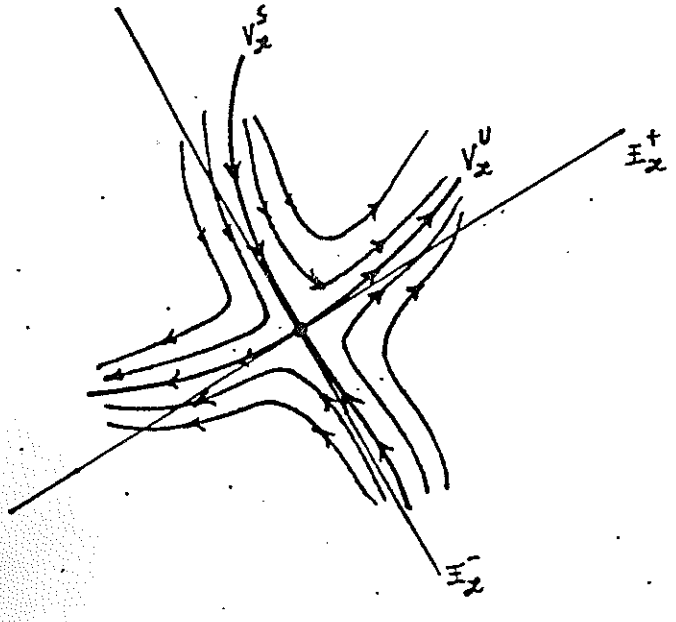


Figure C.57: The figure shows the stable (V^s_x) and unstable (V^u_x) manifold of an equilibrium point x, with the tangent linear subspaces E^-_x and E^+_x.

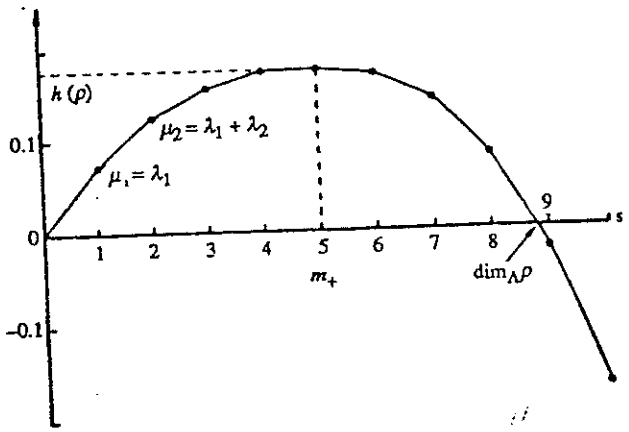


Figure C.58: *Determination of the Liapunov Dimension.* The Liapunov dimension is obtained as the unique root of the function $\Lambda(s)$ ($\Lambda(j)$ is the sum of the Liapunov exponents $\sum_{\lambda_i>\lambda_j} \lambda_i$). The dimension of the unstable manifold is L^+. The graph is from [Ruelle,89] (p. 82).

Bibliography

[Abarbanel et al,91] **H.D. Abarbanel, R. Brown & M.B. Kennel** (1991) "Liapunov Exponents in Chaotic Systems: their Importance and their Evaluation using Observed Data", *Int. J. of Mod. Physics B,* **5**,*9,1347-75.*

[Arnold,80] **V.I. Arnold**, *Chapitres Suplémentaires de la Théorie des Equations Différentielles Ordinaires*, Mir, 1980.

[Atmanspacher et al,88] **H. Atmanspacher, H. Scheingraber & W. Voges** (1988) "Global Scaling Properties of a Chaotic Attractor Reconstructed from Experimental Data", *Phys. Rev.* **A 37**, *1314.*

[Besicovitch,34] **A.S. Besicovitch** (1934) "On the Sum of Digits of Real Numbers Represented in the Dyadic System", *Math. Annalen* **110**, *321-30.*

[Bajo et al, 92] **O. Bajo, F. Fernández & S. Sosvilla** "Volatilidad y Predecibilidad en las Series de Tipo de Cambio Peseta–Dólar: Un Enfoque basado en el Caos Determinista", *Revista Española de Economía, 1992, Monográfico Mercados Financieros, 91-109.*

[Barnsley,88] **M. Barnsley**, *Fractals Everywhere*, Academic Press, 1988.

[Bartlett,90] **M.S. Bartlett** (1990) "Chance or Chaos? (with discussion)", *J. R. Statist. Soc. A,* **153**, *3, 321-347.*

[Ben-Mizrachi et al,84] **A. Ben-Mizrachi, P. Grassberger & I. Proccacia** (1984) "Characterization of experimental (noisy) strange attractors", *Phys. Rev. A (3)* **29**, *975-77.*

[Billingsley,78] **P. Billingsley**, *Ergodic Theory and Information,* Krieger, New York, 1978.

[Billingsley,79] **P. Billingsley**, *Probability and Measure,* Wiley, New York, 1979.

[Box & Jenkins,76] **G.E.P. Box & G.M. Jenkins**, *Time Series Analysis,* Holden-Day, San Francisco, 1976.

[Brandstäter et al,83] **A. Brandstäter, J. Swift, H.J. Swinney, A. Wolf, J.D. Farmer, E. Jen & J.P. Crutchfield** (1983) "Low–Dimensional Chaos in a Hydrodynamical System", *Phys. Rev. Lett.* **51**, *1492-95.*

[Brock,86] **W.A. Brock** (1986) "Distinguishing Random and Deterministic Systems", *Journal of Economic Theory* **40**,*168-195.*

[Brock & Dechert,91] **W.A. Brock & W.D. Dechert**, (1991) " Non-Linear Dynamical Systems: Instability and Chaos in Economics", in *Handbook of Mathematical Economics, vol. IV*, W. Hildebrand and H. Sonnenschein eds., Elsevier Science Publishers BV, 1991.

[Brock et al,87] **W.A. Brock, W.D. Dechert & J.A. Scheinkman** (1987) "A Test for Independence based upon the Correlation Dimension", Department of Economics, University of Wisconsin, Madison, University of Houston and University of Chicago.

[Brock & Sayers,88] **W.A. Brock & C. Sayers** (1988) "Is the Business Cycle Characterized by Deterministic Chaos?", *J. Monetary Econ.* **22**, *71-90.*

[Brock et al,92] **W.A. Brock, D.A. Hsieh & B. LeBaron**, *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*, The MIT Pres, 1992.

[Broomhead & King, 86] **D.S. Broomhead & P. King** (1986) "Extracting Qualitative Dynamics from Experimental Data", *Physica* **20D**, *217-36.*

[Buzug et al,90] **T. Buzug, T. Reimers & G. Pfister** (1990) "Optimal Reconstructions of Strange Attractors from Purely Geometrical Arguments", *Europhys. Lett.,* **13** *(7), 605-10.*

[Canizzo et al,90] **J.K. Canizzo, D.A. Goodings & J.A. Mattei**, (1990) "A Search for Chaotic Behavior in the Light Curves of three Long-Period Variables", *The Astrophysical Journal,* **357**, *235-242.*

[Caswell & Yorke,86] **W.E. Caswell & J.A. Yorke** (1986) "Invisible Errors in Dimension Calculation: Geometrical and Sistematic Effects", in *Dimension and Entropies in Chaotic Dynamical Systems–Quantification of Complex Behavior"*, G. Mayer-Kress ed., (Springer Series in Synergetics, vol. 32) Springer-Verlag, 1986.

[Conde & Calderón,86] **J.J. Conde & J. Calderón** (1986) "Análisis Convencional de un Registro de Oleaje", Programa de Clima Marítimo. Publicación No 10.

[Crutchfield & McNamara,87] **J.S. Crutchfield & B.S. McNamara** (1987) "Equations of Motion from Data Series", *Compl. Syst.* **1**, *417-52.*

[Cutler,90a] **C.D. Cutler** (1990) "Connecting Ergodicity and Dimension in Dynamical Systems", *Ergodic Th. Dynam. Syst.* **10**, *451-62.*

[Cutler,90] **C.D. Cutler** (1990) "A Dynamical System with Integer Information Dimension and Fractal Correlation Exponent", *Comm. Math. Phys.* **129**, *621-29.*

[Cutler,91] **C.D. Cutler** (1991) "Some Results on the Behaviour and Estimation of the Fractal Dimension of Distributions on Attractors", *J. Stat. Phys.*, **62**, *651-708.*

[Deliu et al,91] **A. Deliu, J.S. Geronimo, R. Shonkwiler & D. Hardin** (1992) "Dimensions Associated with Recurrent Self-Similar Sets", *Math. Proc. Camb. Phil. Soc.* **110**, *327-36.*

[Denker & Keller,86] **M. Denker & G. Keller** (1986) "Rigorous Statistical Procedures for Data from Dynamical Systems", *J. Stat. Physics,* 44, *1/2, 67-93.*

[Devaney,87] **R.L. Devaney** *An Introduction to Chaotic Dynamical Systems,* New York, Addison-Wesley, 1987.

[Eckmann et al,86] **J.P. Eckmann, S.O. Kamphorst, D. Ruelle & S. Ciliberto** (1986) "Liapunov Exponents from Time Series", *Physical Review A,* **34**, *6, 4971.*

[Eckmann & Ruelle,85] **J.P. Eckmann & D. Ruelle** (1985) "Ergodic Theory of Chaos and Strange Attractors", *Rev. Mod. Phys.*, **57**, *617.*

[Essex & Nerenberg,91] **C. Essex & M. Nerenberg,** (1991) "Comments on 'Deterministic Chaos: The Science and The Fiction' by D. Ruelle", *Proc. R. Soc. Lond. A* **435**, *287-92.*

[Falconer,90] **K.J. Falconer,** *Fractal Geometry. Mathematical Foundations and Applications,* Wiley, New York, 1990.

[Farmer et al,83] **J.D. Farmer, E. Ott & J.A. Yorke** (1983) "The Dimension of Chaotic Attractors", *Physica* **7D**, *153-80.*

[Fraser,89] **A.M. Fraser** (1989) "Reconstructing Attractors from Scalar Time Series: a Comparison of Singular System and Redundancy Criteria", *Physica* **34D**, *391-404.*

[Fraser & Swinney,86] **A.M. Fraser & H.L. Swinney** (1986) "Independent Coordinates for Strange Attractors from Mutual Information", *Phys. Rev. A* **33**, *2, 1134-40.*

[Frederickson et al,83] **P. Frederickson, J.L. Kaplan, E.D. Yorke & J.A. Yorke** (1983) "The Liapunov Dimension of Strange attractors", *J. Diff. Equ.*, **49**, *185.*

[Froehling et al,81] **H. Froehling, J. Crutchfield, J.D. Farmer, N. Packard & R. Shaw** (1981) "On Determining the Dimension of Chaotic Flows", *Physica D* **3**, *605.*

[Gantmacher,77] **F.R. Gantmacher**, *Matrix Theory*, Chelsea, 1977.

[Guillemin & Pollac,74] **V. Guillemin & A. Pollac**, *Differential Topology*, Prentice Hall, 1974.

[Gonin & Money,89] **R. Gonin & A.H. Money** (1989) *Non-Linear L_p-Norm Estimation*, Marcel Dekker, Inc. Statistics Texbooks and Monographs, vol. 100.

[Grassberger,83] **P. Grassberger** (1983) "Generalized Dimensions of Strange Attractors", *Phys. Lett. A*, **97**, *227.*

[Grassberger,88] **P. Grassberger**, (1988) "Finite Sample Corrections to Entropy and Dimension Estimates", *Physics Letters A* **128**, *369-73.*

[Grassberger,90] **P. Grassberger**(1990) "An Optimized Box–Assisted Algorithm for Fractal Dimensions", *Phys. Lett. A*, **148**, *63.*

[Grassberger et al,91] **P. Grassberger, T. Schreiber & C. Schaffrath** (1991) "Nonlinear Time Sequence Analysis", *Int. J. Bif. Chaos*, **1**, *3, 521-547.*

[Grassberger & Proccacia,83a] **P. Grassberger & I. Procaccia** (1983a) "Characterization of Strange Attractors", *Phys. Rev. Lett.*, **50**, *346-49.*

[Grassberger & Proccacia,83b] **P. Grassberger & I. Procaccia** (1983) "Estimation of the Kolmogorov Entropy from a Chaotic Signal", *Phys. Rev. A*, **28**, *2591-93.*

[Grassberger & Procaccia,83c] **P. Grassberger & I. Procaccia** (1983c) "Measuring the Strangeness of Strange Attractors", *Physica D*, **9**, *189-208.*

[Greenside et al,82] **H. Greenside, A. Wolf, J. Swift & T. Pignataro** (1982) "Impracticality of a Box Counting Algorithm for Calculating the Dimensionality of Strange Attractors", *Phys. Rev. A*25, *3453.*

[Haken,83] **H. Haken** (1983) "At Least One Liapunov Exponent Vanishes if the Trayectory of an Attractor does not Contain a Fixed Point", *Physics Letters A,* **94**, *2, 71-72.*

[Halsey et al,86] **T.C. Halsey, M.H. Jensen, L.P. Kadanoff, I. Procaccia & B.I. Shraiman** (1986) "Fractal Measures and their Singularities: the Characterization of Strange Sets", *Phys. Rev. A,* **33**, *1141-51.*

[Henon,76] **M. Henon** (1976) "A Two–Dimensional Mapping with a Strange Attractor",*Comm. Math. Physics* **50**, *69.*

[Holden & Muhamad, 86] **A.V. Holden & M.A. Muhamad** (1986) "A Graphical Zoo of Strange and Peculiar Attractors", in *Chaos*, A.V. Holden ed., Princeton Univ. Press.

[Holzfuss & Mayer-Kress,86] **J. Holzfuss & G. Mayer–Kress** (1986) "An Approach to Error–Estimation in the Application of Dimension Algorithms", in *Dimension and Entropies in Chaotic Dynamical Systems–Quantification of Complex Behavior"*, G. Mayer-Kress ed., (Springer Series in Synergetics, vol. 32) Springer-Verlag, 1986.

[Horn & Johnson,85] **R.A. Horn & C.R. Johnson** , *Matrix Analysis*, Cambridge, 1985.

[Hunt et al] **B. Hunt, T. Sauer & J.A. Yorke** "Prevalence: a Traslation–Invariant 'Almost Every' on Infinite-Dimensional Spaces", preprint.

[Ilyachenko,83] **Y.S. Ilyachenko** (1983) "On the Dimension of Attractors of k-Contracting Systems in an Infinite-Dimensional space", *Vestn. Mosk. Univ. Ser. 1 Mekh,* **3**, *52-58.*

[Irwin,80] **M.C. Irwin**, *Smooth Dynamical Systems,* Academic Press, 1980.

[Isham,93] **V. Isham** (1993) "Statistical Aspects of Chaos: A Review" in *Network and Chaos–Statistical and Probabilistic Aspects*, O.E. Barndorff-Nielsen, J.L. Jensen and W.S. Kendall eds., Chapman & Hall, 1993.

[Jensen,93] **J.L. Jensen** (1993) "Chaotic Dynamical Systems with a View towards Statistics: A Review", in *Networks and Chaos–Statistical and Probabilistic Aspects*, O.E. Barndorff-Nielsen, J.L. Jensen and W.S. Kendall eds., Chapman & Hall, 1993.

[Johnson et al,87] **R.A. Johnson, K.J. Palmer & R. Sell** (1987) "Ergodic Properties of Linear Dynamical Systems", *SIAM J. Math. An.*, **18**, *1-33.*

[Kaplan & Yorke,78] **J. Kaplan & J.A. Yorke** (1978) "Functional Differential Equations and the Approximation of Fixed Points", Lecture Notes in Math. 730, 228 (Peitgen & Walters eds.), Springer.

[Koopmans,74] **L.H. Koopmans**, *The Spectral Analysis of Time Series*, Academic Press, New York, 1974.

[Krengel,85] **U. Krengel** *Ergodic Theorems,* De Gruyter Studies in Mathematics, W. de Gruyter, 1985.

[Ledrappier & Young, 84] **F. Ledrappier & L.S. Young** (1984) "The Metric Entropy of Diffeormorphisms", *Bull. Am. Math. Soc. (New Ser.)* **11**, *343.*

[Liebert et al,91] **W. Liebert, K. Pawelzik & H.G. Schuster** (1991) "Optimal Embeddings of Chaotic Attractors from Topological Considerations", *Europhys. Lett.* **14** *(6), 521-26.*

[Liebert & Schuster,89] **W. Liebert & H.G. Schuster** (1989) "Proper Choice of the Time Delay for the Analysis of Chaotic Time Series", *Phis. Lett. A,* **142***(2,3), 107-11.*

[Lorenz,63] **E.N. Lorenz** (1963) "Deterministic Nonperiodic Flow", *J. Atmos. Sci.* **20**, *130.*

[Lorenz,90] **E.N. Lorenz**, (1990) "Dimension of Weather and Climate Attractors", *Nature* **353**, *241-44.*

[Mattila,75] **P. Mattila** (1975) "Hausdorff Dimension, Orthogonal Projections and Intersection with Planes", *Ann. Acad. Sci. Fenn. Ser. A1, Math.*, **1**, *227.*

[Mañé,81] **R. Mañé**, *Ergodic Theory and Differentiable Dynamics,* Springer-Verlag, 1987.

[Mayer-Kress,87] **G. Mayer–Kress** (1987) "Application of Dimension Algorithms to Experimental Chaos" in *Chaos*, Hao Bai-Lin ed., World Scientfic, 1987.

[Mera et al,93] **E. Mera, M. Morán & J.M. Rey**, (1993) "Cálculo de la Dimension Box–Counting en Bajas Dimensiones", unpublished.

[Mera & Morán] **E. Mera & M. Morán** "Convergence of the Eckmann–Ruelle algorithm for computing Liapunov Exponents", manuscript in preparation.

[Morán & Rey,94] **M. Morán & J.M. Rey** (1994) "Singularity of Self–Similar Measures with respect to Hausdorff Measures", preprint.

[Osborne et al,86] **A.R. Osborne, A.D. Kirwan, A. Provenzale & L. Bergamasco** (1986) "A Search for Chaotic Behavior in Large and Mesoscale Motions in the Pacific Ocean", *Physica D* **35**, *357.*

[Osborne & Provenzale,89] **A.R. Osborne & A. Provenzale** (1989) "Finite Correlation Dimension for Stochastic Systems with Power Law Spectra", *Physica D*, **35**, *357-381.*

[Oseledec,68] **V.I. Oseledec** (1968) "A Multiplicative Ergodic Theorem. Liapunov Characteristic Exponents", *Trans. Moscow Math. Soc.*, **19**, *179-231.*

[Packard et al,1980] **N.H. Packard, J.P. Crutchfield, J.D. Farmer & R.S. Shaw** (1980) "Geometry from a Time Series", *Phys. Rev. Lett.* **45**, *712-716.*

[Priestley,81] **M.B. Priestley**, *Spectral Analysis and Time Series*, vols. I and II, Academic, London, 1987.

[Provenzale & Osborne,89] **A. Provenzale & A.R. Osborne** (1989) "Finite Correlation Dimension for Stochastic Systems with Power-Law spectra", *Physica D*, **35**, *357-381.*

[Provenzale et al,91] **A. Provenzale, A.R. Osborne & R. Soj** (1991) "Convergence of the K_2–Entropy for Random Noises with Power Law Spectra",*Physica D*, **47**, *361-72.*

[Provenzale et al,92] **A. Provenzale, L.A. Smith, R. Vio & G. Murante** (1992) "Distinguishing between Low-dimensional Dynamics and Randomness in Measured Time Series", preprint.

[JRSS,92] Royal Statistical Society Meeting on Chaos (Discussion), (1992) *J. Roy. Stat. Soc. B* **54**, *451-474.*

[Raghunathan,78] **M.S. Raghunathan** (1978) "A Proof of Oseledec's Multiplicative Ergodic Theorem", *Israel J. Math.*, **32**, *4, 356-62.*

[Ramsey & Yuan,89] **J.B. Ramsey & H.-J. Yuan** (1990) "The statistical Properties of Dimension Calculation using Small Data Sets", *Nonlinearity* **3**, *155-176.*

[Renyi,70] **A. Renyi**, *Probability Theory*, North-Holland, Amsterdam, 1970.

[Rössler,76] **O.E. Rössler** (1976) "An Equation for Continuous Chaos", *Phys. Lett. A* **54**, *397.*

[Ruelle,79] **D. Ruelle** (1979) "Ergodic Theory of Differentiable Dynamical Systems", *Publ. Math. IHES,* **50**, *275-307.*

[Ruelle,93] **D. Ruelle**, *Chance and Chaos*, Penguin Books, 1993.

[Ruelle,87] **D. Ruelle**, *Chaotic Evolution and Strange Attractors.*, Cambridge University Press, 1987.

[Ruelle,90] **D. Ruelle** (1990) "Deterministic Chaos: the Science and the Fiction", *Proc. R. Soc. London A,* **427**, *241.*

[Ruelle & Takens,71] **D. Ruelle & F. Takens** (1971) "On the Nature of Turbulence", *Comm. Math. Physics.,* **21**, *12.*

[Russell et al,80] **D.A. Russell, J.D. Hansen & E. Ott** (1980) "Dimensionality and Liapunov Numbers of Strange Attractors", *Phys. Rev. Lett* **45**, *1175.*

[Sakai & Tokumaru,80] **H. Sakai & H. Tokumaru** (1980) "Autocorrelations of a Certain Chaos", *IEEE Transactions on Acoustic, Speech and Signal Processing V.I. ASSP-28, 5, 588-90.*

[Sano & Sawada,85] **M. Sano & Y. Sawada** (1985) "Measurement of the Liapunov Spectrum from a Chaotic Time Series", *Phys. Rev. Lett.,* **55**, *10, 1082-85.*

[Sato et al,87] **S. Sato, M. Sano & Y. Sawada** (1987) "Practical Methods of Measuring the Generalized Dimension and the Largest Liapunov Exponent in High Dimensional Chaotic Systems", *Prog. Theor. Phys.,* **77**, *1.*

[Sauer et al,91] **T. Sauer, J.A. Yorke, & M. Casdagli** (1991) "Embedology" *J. Stat. Physics,* **65**, *579-616.*

[Scheinkman & Lebaron,89] **J. Scheinkman & B. Lebaron** (1989) "Nonlinear Dynamics and Stock Returns", *J.Business,* **62**, *311.*

[Takens,81] **F. Takens** (1981) "Detecting Strange Attractors in Turbulence", *Dynamical Systems and Turbulence* (Rand & Young eds.) Lecture Notes in Mathematics 898, 366 (Springer).

[Takens,83] **F. Takens** (1983) "Invariants related to Dimension and Entropy", in *Atas do 13. Col. Brasiliero de Matematicas, Rio de Janeiro.*

[Theiler,86] **J. Theiler** (1986) "Spurious Dimension from Correlation Algorithms applied to Limited Time-Series Data", *Pys. Rev. A* **34**, *3, 2427-32.*

[Theiler,90] **J. Theiler** (1990) "Estimating Fractal Dimension", *J. Opt. Soc. Am. A*, **7**, *1055-73.*

[Theiler,90b] **J. Theiler** (1990) "Statistical Precision of Dimension Estimators", *Phys. Rev. A* **41**, *6, 3038-51.*

[Theiler,91] **J. Theiler** (1991) "Somme Comments on the Correlation Dimension of $1/f^{\alpha}$ noise", *Phys. Lett. A*, **155**, *480.*

[Turbulence Seminar,75] **P. Bernard & T. Ratiu (eds.)**, *Turbulence Seminar*, Proceedings 1976/77, Lecture Notes in Mathematics 615, Springer-Verlag.

[Tong,90] **H. Tong**, *Non-Linear Time Series Analysis*, Oxford Univ. Press, 1990.

[Tricot,82] **C. Tricot** (1982) "Two definitions of Fractional Dimension" *Proc. Camb. Phil. Soc.*, **91**, *57-74.*

[Vastano & Kostelich,85] **J.A. Vastano & E.J. Kostelich** (1985) "Comparison of Algorithms for Determining the Liapunov Exponents from Experimental Data", in *Dimensions and Entropies in Chaotic Systems,* G. Mayer–Kress, ed. (Springer).

[Walters,82] **P. Walters**, *An Introduction to Ergodic Theory*, Springer–Verlag, 1982.

[Wolf et al,84] **A. Wolf, J.B. Swift, H.L. Swinney & A. Vastano** (1984) "Determining Liapunov Exponents from Time Series", *Physica D*, **16**, *285-317.*

[Wolff,92] **R. Wolff** (1992) "Local Liapunov Exponents: Looking Closely at Chaos", *J.R. Stat. Soc. B* **54**, *353-71.*

[Young,82] **L.-S. Young** (1982) "Dimension, Entropy and Liapunov Exponents", *Ergod. Theor. & Dyn. Sys.*, **2**, *109.*